DISSERTATIONES
SOCIOLOGICAE
UNIVERSITATIS
TARTUENSIS
23

KRISTJAN KIKERPILL

Crime-as-communication: detecting diagnostically useful information from the content and context of social engineering attacks

TARTU ÜLIKOOL
UNIVERSITAS TARTUENSIS
1632

DISSERTATIONES SOCIOLOGICAE UNIVERSITATIS TARTUENSIS

23

DISSERTATIONES SOCIOLOGICAE UNIVERSITATIS TARTUENSIS

23

KRISTJAN KIKERPILL

Crime-as-communication: detecting diagnostically useful information from the content and context of social engineering attacks

UNIVERSITY OF TARTU
Press

Institute of Social Studies, University of Tartu

Dissertation accepted in fulfilment of the requirements for the degree of Doctor of Philosophy (in Sociology) on September 23, 2021, by the Council of the Institute of Social Studies, University of Tartu.

Supervisor:	Professor Andra Siibak Institute of Social Studies, University of Tartu
Opponent:	Dr Elisabeth Carter University of Roehampton, the United Kingdom
Commencement:	February 1, 2022, at the University of Tartu

The publication of this dissertation is granted by the Institute of Social Studies, University of Tartu. Part of the research in this dissertation was supported by a personal scholarship awarded to the author by the Estonian Police and Border Guard Board.

ISSN 1736-0307
ISBN 978-9949-03-740-7 (print)
ISBN 978-9949-03-741-4 (pdf)

University of Tartu Press
www.tyk.ee

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

Study I: Kikerpill, K. & Siibak, A. 2019. Living in a Spamster's Paradise: Deceit and Threats in Phishing Emails. *Masaryk University Journal of Law and Technology*, *13*(1): 45–63. DOI: 10.5817/MUJLT2019-1-3

Study II: Kikerpill, K. & Siibak, A. 2021. Abusing the COVID-19 Pan(de)mic: A Perfect Storm for Online Scams. In J. C. Pollock & D. A. Kovach (Eds.) *COVID-19 in International Media: Global Pandemic Perspectives*. New York: Routledge. DOI: 10.4324/9781003181705-25

Study III: Kikerpill, K. & Siibak, A. 2021. Mazephishing: The COVID-19 pandemic as credible social context for social engineering attacks. *Trames Journal of the Humanities and Social Sciences* (upcoming).

Study IV: Kikerpill, K. 2021. The Individual's Role in Cybercrime Prevention: Internal Spheres of Protection and Our Ability to Safeguard Them. *Kybernetes*, *50*(4), 1015–1026. DOI: 10.1108/K-06-2020-0335

AUTHOR'S CONTRIBUTION

Study I. The author was solely responsible for data collection, data analysis and writing the article. The second author contributed to editing the manuscript.

Study II. The author was solely responsible for data collection and writing the chapter. The second author contributed to data analysis and editing the manuscript.

Study III. The author was solely responsible for data collection and mainly responsible for data analysis and writing the article. The second author wrote minor parts and edited the manuscript.

Study IV. The author was solely responsible for researching and writing the article.

ACKNOWLEDGEMENTS

At its core, I believe the academic journey to be a forgiving one: everyone is just tall enough to concurrently have their head in the clouds and their feet firmly on the ground. Along the way, a whole group of people are there to help you avoid and recover from unexpected bumps as well as to prevent unwarranted levitation.

Firstly, I want to give my deepest thanks to my supervisor, Prof. Andra Siibak. This thesis would not have been possible without Andra’s relentless support, encouragement, and willingness to read essay-length e-mails. Thank you!

I am grateful to Associate Prof. Maria Murumaa-Mengel, Prof. Veronika Kalmus, Simone Eelmaa and Taivo Liivak. Maria’s unending energy is a sight to behold and a phenomenon to learn from. Thank you for involving me in all kinds of interesting stuff. Prof. Kalmus taught me a great deal about giving and receiving feedback. I want to thank Simone for her impressive determination, and for our discussions regarding “legally undercounselled people”. I am grateful to Taivo for helping me help myself out when times were dire. My thanks also go to the PhD committee for their constructive criticism and for showing me that a lot could grow from this thesis in the future. I am also thankful to the Estonian Police and Border Guard Board for awarding me a scholarship that contributed time and space for finishing my thesis.

I would be remiss if I did not give thanks to my friends Maria and John. Granted, when our names are spoken in the same sentence, it may lead people to expect that a dark, religion-themed comedy is afoot. They would not be wrong.

My warmest thanks go to my mom, Merike, and my sister, Mari. The strength and resilience of my mother makes documentaries about survival seem like sitcoms. My sister has always provided the oh-so-necessary element of sibling rivalry. I am also forever grateful to the Tartu chapter of my family: my grandmother and grandfather, and Merle and Gunnar. The wisdom and humanity of my grandmother, Hilda, is unmatched. The patience and knowledge of my grandfather, Juhan, is the inspiration for my own knowledge-seeking endeavours.

INTRODUCTION

Modern information and communication technologies (ICTs) have created new ways for doing old things. Working, playing, shopping, participating in events and discussions are increasingly reliant on the transmission and receipt of data. It is possible to note that regardless of the verbs we use in describing our "on-line" activities, such activities are all inherently rooted in and dependent upon communication. Consequently, this means that violating social norms, including committing crimes, has also been adapted to the newer way of being. Since crime is a socially constructed phenomenon (Posick, 2018), it goes where people go.

While modern communication technology has made it possible to easily contact people regardless of their physical location, the use of such technologies and the access to vast amounts of information increasingly causes people to experience the effects of information overload (Gunaratne et al., 2020), i.e. an adverse state in which a decision-maker's usual cognitive abilities are hindered. Hence, the previously described situation is exacerbated by the fact that some, who make use of the same possibilities afforded by modern technology and the easy access it provides, do so for the purposes of committing crimes (**Study IV**), i.e. cybercrimes (see also McGuire & Dowling, 2013). In fact, the rates of globally perpetrated cybercriminal acts have been rapidly increasing (PurpleSec, 2021) and show no signs of slowing down.

According to a recent report (Proofpoint, 2019: 19), 99% of cybercrime threats require some human interaction – opening a file, following a link or opening a document – by the recipient to be successful. This means that the malicious use of social engineering, which is defined in modern security discussions as acts that influence a person to take an action that is not in their best interests (see 2.3 below; Hadnagy, 2018; Hatfield, 2018), is key to criminals' success (Proofpoint, 2019).

Acknowledging the role played by social engineering and, in particular, influencing in the growing problem of cybercrime is to also recognise that prevention efforts must be focussed on a specific moment in time, i.e. the brief period of time between a person receiving a message and taking action based on their interpretation of the content received (see e.g. McAlaney & Hills, 2020). Since people's awareness of the cybercrime problem is increasing concurrently with a decrease of confidence in being able to stay safe online (European Commission, 2020) and the applicability of current criminological theories to the problem of cybercrime are considered inconclusive at best (Button & Cross, 2017; **Study IV**), we need a new approach to understanding technology-mediated crime.

In my thesis, I argue that this approach, which I call **crime-as-communication**, can be provided by combining aspects of environmental criminology (how criminals and crime targets converge in space and time: Felson & Cohen, 1979) with ideas from the disciplines of law (how a crime target's "will to act" is envisioned and understood in criminal offences such as fraud and extortion), communication (how meaning is encoded, transmitted and understood: Hall, 1973; Levine, 2019) and social psychology (what techniques are used to influence people to gain their compliance in crimes requiring action by the victim: Cialdini, 2009) as well as media sociology and the sociology of deviance.

The prevention of victimisation that entails from crime acts committed as communicative acts, which rely on deception and manipulated messages-in-context, can only occur if the deception embedded in the message that is further amplified by its context is detected. According to Levine (2019), people, by default, assume that communication is truthful, because this disposition helps us function in a social world. Because intentional deception is the exception not the rule (Levine, 2019), people are also more vulnerable to instances where deceptive practices are in fact employed by those producing the messages (Hall, 1973). In order for suspicion to arise about the content of certain acts of communication, there must be a triggering event, i.e. something in the message must constitute 'diagnostically useful information' (Levine, 2019) that triggers a person to suspect deception (and, consequently, the potentially criminal nature of the communication).

Recipients' detection of deception carries significance for another reason: the fact of just how alone we are in that brief moment between receiving a message and deciding what to do next. Suggesting that third persons, e.g. parents, colleagues or law enforcement officers, are able to intervene on behalf of the recipient, i.e. act as capable guardians against crime (Felson & Cohen, 1979), would require said third persons to have real-time access to our communications (**Study IV**). Thus, the protective role of others, e.g. parents protecting their children or law enforcement officers protecting the public, is severely diminished in the particular moment of the communicative act. In other words, knowledge about the mechanics and inner workings of scams is crucial.

The aforementioned knowledge can be fostered preventively by parents (Smahel et al., 2020), places of employment (MacEwan, 2017; Nguyen, Jensen, Durcikova, & Wright, 2020) or public awareness campaigns (Sasse & Smith, 2016; Button & Cross, 2017; Whitty, 2019), or learned "the hard way" after the fact (Button & Cross, 2017). Nevertheless, due to the rapid nature of technology-mediated communications, the target of a crime (the recipient) can only rely on their own current knowledge regarding the initial interpretation of incoming (crime-as-)communication, because asking for help or seeking verification from external sources already assumes that something in the original communication triggered suspicion in the recipient.

Although susceptibility to social engineering attacks has become a popular research topic (see Nguyen, Jensen, Durcikova, & Wright, 2020: Appendix A), research about the content of social engineering attacks, e.g. which persuasive strategies are used by criminals when crafting socially engineered messages and how these appear in the messages (Kim & Kim, 2013; Wright, Jensen, Thatcher, Dinger, & Marett, 2014; Carter, 2015; Zielinska, Welk, Mayhorn, & Murphy-Hill, 2016; Stojnic, Vatsalan, & Arachchilage, 2021), is still lacking.

This lack of focus on persuasive content itself limits our understanding on what constitutes relevant diagnostically useful information. Even less is known about the influence that immediate social context has on the ways in which perpetrators produce (Verma, Crane, & Gnawalli, 2018; Steinmetz, Pimentel, & Roe, 2021) and recipients interpret (Greene, Steves, Theofanos, & Kostic, 2018) the content of phishing attacks. Phishing attacks are social engineering attacks commonly perpetrated via email (Khonji, Iraqi, & Jones, 2013), but also through text messages (SMS phishing or smishing) and phone calls (voice phishing or vishing) (Hong, 2012).

Hence, if the key to criminals' success in 99% of cybercrimes is deceiving recipients into doing their bidding by using persuasive messages(-in-context), then the solution to preventing 99% of cybercrimes must start with providing recipients with the knowledge necessary for detecting the aforementioned deception.

Thus, **the aim of my thesis is to explain the role and importance of interpreting messages-in-context to distinguish potentially criminal input from all received input for the purpose of preventing victimisation from social engineering attacks.**

The thesis is based on four articles, of which three are empirical studies (**Study I, Study II, Study III**) and one (**Study IV**) presents a theoretical criticism of current criminological thought. The arguments I present in my thesis operate in an international context and address universal phenomena, i.e. the use, and the interpretation of the content, of deceptive communication in social engineering attacks.

The cover text is structured as follows. Firstly, I will set the problem and provide reasons why there is an urgent need to better our understanding of mediated crimes, including cybercrimes, in particular with respect to the changing responsibilities of and increased expectations placed upon individuals. In Chapter 2, I provide the theoretical background which lead me from the study of environmental criminology proper to an approach that is motivated by environmental criminology but situated within a communication-based framework and present my research questions. In Chapter 3, I describe the methods used in the studies that form the basis for this cover text, and in Chapter 4, I provide the results and analysis of the results of the previously mentioned empirical studies. Chapter 5 provides a discussion of the results as answers to my research questions and further motivations for and benefits of adopting the **crime-as-communication** approach to the study of mediated and cybercrimes.

1. SETTING THE PROBLEM

The core problem under discussion in my thesis is how insights from communication can inform, instruct and potentially mitigate the negative outcomes from cybercriminal acts targeting individuals. However, and as with any academic endeavour, the ideas that ultimately shape the arguments presented in the following text are framed and influenced by what has been done before. For instance, while the core of my arguments emerge from a synthesis of opportunity theories of crime (see Chapter 2.1) and scholarship on deceptive communication in the form of social engineering (see Chapter 2.3), these approaches function well on a small scale, i.e. on the level of the crime act itself. Although useful in the role I have assigned them, the previously mentioned approaches require additional help from concepts that aid in framing the importance of the work I present. Thus, before diving the reader into the minutia of criminological thought as well as the historical and current understanding of 'social engineering', I am obligated to explain what lead me to choose such theoretical approaches and not others.

The point of departure may seem like an odd one: the adoption and entry into force of Directive (EU) 2016/1148 concerning measures for a high common level of security of network and information systems across the (European) Union (Directive (EU) 2016/1148). This piece of legislation is also referred to as "the NIS Directive" (network and information security) (Markopoulou, Papakonstantinou, & De Hert, 2019), and at this point, should further be understood as "the original" version, given that the Commission submitted a proposal for version 2.0 in late 2020. The significance of the NIS Directive came, primarily, from it being the first legislative proposal aimed at creating a Union-wide approach to cybersecurity. The EU defines cybersecurity as "activities necessary to protect network and information systems, the users of such systems, and other persons affected by cyber threats" (Regulation (EU) 2019/881).

In two official documents, an impact assessment accompanying the proposal for the original NIS Directive (European Commission, 2013) and a Joint Communication following the adoption of the Directive (European Commission, 2017), the issue of "human error" as a problem for cybersecurity is raised. In the impact assessment (European Commission, 2013), human error is included alongside a variety of factors, e.g. natural events, technical failures and malicious attacks, contributing to the occurrence of cyber incidents (or accidents). In the Joint Communication (European Commission, 2017), the approach to human error became more specific, as the document cited a survey by IBM Global Technology Services (2014), which suggested that 95% of (cyber)-security incidents recognise "human error" as a contributing factor. As noted above, this number has now increased to 99% (Proofpoint, 2019). Furthermore, in its Cybersecurity Act (Regulation (EU) 2019/881: recital 8), the European Union also institutionalised the concept of cyber-hygiene, i.e. "simple, routine measures that, where implemented and carried out regularly by citizens, organisations and businesses, minimise their exposure to risks from cyber threats".

Thus, an issue that a union of sovereign states deems inescapably important lead that same union to compel private companies to implement "appropriate and proportionate technical and organisational measures to manage the risks posed to the security of networks and information systems" (Directive (EU) 2016/1148: art 14). Since companies are legal abstractions (subjects of law established pursuant to law, General Part of the Civil Code Act, § 24; see also Naffine, 2003), i.e. companies can be established, dissolved, sold, acquired, merged and divided but you still cannot point your finger at a "company", the responsibility for carrying out those security measures is placed on employees, managers and executives. In other words, the people considered to be a contributing factor to 95–99% of security incidents (IBM Global Technology Services, 2014; Proofpoint, 2019) became the vehicles for organisational and, implicitly, state-level cybersecurity.

This notion, i.e. a state encouraging or compelling individuals to acknowledge and assume a degree of responsibility for managing their own risks (Burchell, 1996: 6), is known as "responsibilisation". Burchell (1996) and Garland (1996) were the first to use this concept in the social sciences (referred to in Brown, 2021). Garland (1996), in particular, applied the concept of responsibilisation to the issue of crime control. Garland viewed this responsibilisation strategy as the central government's effort to "act upon crime not in a direct fashion through state agencies" (1996: 452), e.g. police, courts, prisons or social work, but instead operate indirectly and activate private actors and persuade them to act appropriately.

As mentioned earlier, while human error was highlighted as a contributing factor to security incidents, malicious attacks were also included in the discussion of threats (European Commission, 2013). Broadly taken, threats to organisational cybersecurity come from 1) malicious insiders, 2) outsiders, 3) a combination of 1 and 2 or 4) inadvertent actors (IBM Global Technology Services, 2014). Aside from inadvertent actors, the remaining sources of threats certainly call into consideration the possibility of criminal activity.

Regarding the previously mentioned general threat distribution, more than 50% of IT decision-makers (PurpleSec, 2021) claim that the main threat to their organisation is phishing. Furthermore, a recent survey assessed 81% of employees to be "risky" when it comes to detecting phishing attacks (MediaPro, 2020). Moreover, when employees are forced into a remote work situation, e.g. during the COVID-19 pandemic, employees themselves admit to increasing their employer's cybersecurity risks (Help Net Security, 2021). Hence, we are at an impasse – the state compels private companies to adopt appropriate technical and organisational cybersecurity measures; in practice, carrying out such tasks falls on employees, managers and executives, i.e. people, and the personnel is consistently considered as "risky" when it comes to the primary threat, i.e. phishing attacks, that companies face.

Since phishing attacks, which fall under the category of “outsider threats”, are also crimes (**Study I**), it means that employees are also involved in, and essentially tasked with, crime control. While most of us would probably agree that it is a good idea to lock your door at night, it is not an obligation enforced by the possibility of sanctions. Being taught or asked to lock the door is vastly different from the state obligating you to do so, even if implicitly and in a “trickle-down” fashion (**Study IV**). Furthermore, the concept of responsibilisation has helped to highlight the problem with respect to employees working for companies subject to relevant security obligations. It does not, however, even begin to address the larger social problem regarding people whose awareness of the cybercrime problem is increasing but whose confidence in their ability to stay safe online is decreasing (European Commission, 2020). The previously described issue is exacerbated by the fact that even law enforcement rarely knows the background of the criminals perpetrating deception-based cyber-attacks (Button, Lewis, & Tapley, 2009: 13), no one demographic is considered more or less vulnerable to such deceptive acts, and we lack a consensus on the underlying psychological reasons for why people comply with mediated fraudulent requests (Norris, Brookes, & Dowell, 2019).

Nevertheless, because each phishing attack is a crime or an attempt at one (**Study I**), we can start untangling this social problem by focussing on phishing attacks as specific crime acts. In current criminological thought, opportunity theories of crime stand out due to their focus on the crime act, not the criminals (Clarke, 2013), and thus provide a suitable point of departure. While phishing attacks are crimes or crime attempts, they are, at the same time, communicated acts of social engineering, i.e. acts that influence “a person to take an action that may or may not be in their best interests” (Hadnagy, 2018: 7). Therefore, to understand the threat of phishing attacks, it is important to understand both 1) the crime act and 2) the communicative act. In essence, we need to understand the violation of social norms in the form of criminal activity (as opposed to lawful activities) and deceptive communication (as opposed to truthful communication).

2. THEORETICAL BACKGROUND

2.1 Opportunity theories of crime: the routine activities approach

Although wrought with institutional resistance and initial struggles for acceptance (Felson, 2008; Felson & Clarke, 2011), opportunity theories of crime have now become the mainstay of criminological thought (Wilcox & Cullen, 2018). According to Brantingham and Brantingham (1981), there are four dimensions to understanding crime: the legal dimension, the offender dimension, the victim dimension and the place (or situational) dimension. As opposed to traditional criminology (see Felson, 2019: 612; Andresen, 2010: 6–7), which focusses primarily on the first three dimensions and seeks to understand criminal motivation and criminality, the environmental criminology approach concerns itself with the fourth dimension of place and/or the crime situation. The term "environmental criminology" was coined by Jeffery in his book *Crime Prevention Through Environmental Design* (1971, as cited in Andresen, 2010: 6) as a call for the establishment of a new school of thought in criminology.

The first complete approach published as a journal article or book chapter in the environmental criminology literature was the routine activity approach by Felson & Cohen (1979; see also Andresen, 2010: 13–15). Provided that the routine activity approach was a clear break with ideas present in conventional criminological thought at the time, in particular with social disorganization theory (Andresen, 2010), Felson has described in detail the institutional resistance the approach met at first and how difficult it was to get the original work published (2008). This struggle is not unique and was similarly experienced by other proponents of the environmental criminology approach (see the historical account by Ronald Clarke in Felson & Clarke, 2011).

The routine activity approach was sourced from Amos Hawley's (1950) study of human ecology and focussed on the daily, regular activities of individuals (as distributed in time and space) as these routines and activities produced opportunities for predatory criminal exploitation. In their foundational work, Felson & Cohen (1979) analyse how activities away from the household in a post-World War II United States contributed to increases in burglaries. In other words, the way in which people's routine activities changed directly contributed to increases in criminal opportunity even though quality of life was on the rise otherwise.

The basic premise of the routine activities approach posits that most criminal acts "require convergence in space and time of *likely offenders*, *suitable targets* and the *absence of capable guardians* against crime" (emphasis in original) (Felson & Cohen, 1979: 588). Felson (2008) has explained the choice of 'target' instead of 'victim' to come from the fact that where criminals are after material possessions, e.g. valuables, the target and victim are not the same. Put differently, when a burglar takes a TV from its owner, the owner is the victim but the TV was the actual target from the offender's perspective (Felson & Clarke, 1998). Thus, if a likely offender and a suitable target, including where the target is a person, were to converge in space and time, crime is often the result. Hence, the capable guardian is the "element" whose absence allows for crime and whose presence can foil it. Over time, the theory has of course evolved with new elements such as handlers and place managers being added (see e.g. Eck, 2003), but for better or worse, the combination of the three key elements of likely offenders, suitable targets and capable guardians is the most pronounced contribution associated with the routine activity approach (see Felson's own account about such concerns in Felson, 2008).

Since its first publication more than 40 years ago, the routine activity approach has been extensively applied in criminological research regarding so-called 'terrestrial' crime (Spano & Freilich, 2009; Wilcox & Cullen, 2018; see also Perkins, Howell, Dodge, Burruss, & Maimon, 2020: 3–4), i.e. crime without a pronounced technology-mediated element that occurs, as Pease (2001) memorably dubbed it, in "meatspace". When issues of computer-related crime first emerged into criminological inquiry in the early 2000s (Grabosky, 2001; Capeller, 2001), the debate focussed on whether existing criminological theory is sufficient for explaining and predicting criminal acts in the new environment, i.e. "cyberspace". Given the dominance of the routine activity approach, it is somewhat unsurprising that Grabosky (2001) singled it out as a way of explaining the emerging problem of computer-related offences. Even so, a full account of how the routine activity approach may be applied to technology-mediated crime, i.e. how the now well-known constitutive elements may be transposed to explain criminal acts occurring in 'cyberspace', was proposed slightly later by Yar (2005). In fact, Yar's account remains the oft-cited source for transposing the routine activities into the technology-mediated environment and was recently revisited by Leukfeldt & Yar (2016).

Although the routine activity approach has, by now, also seen ample application in the realm of cybercrime (Holt & Bossler, 2008; Bossler & Holt, 2009; Leukfeldt & Yar, 2016), the results are often mixed (Leukfeldt & Yar, 2016) with respect to which elements of the approach can explain the occurrence of cybercrimes. In particular, this concerns the element of the capable guardian or guardianship in general, as the absence of this element suggests that a crime act will be successful (Felson & Cohen, 1979).

Lack of clarity on what exactly constitutes 'capable guardianship' is a core issue of the routine activities approach – cybercrime or otherwise – because it is incredibly easy to keep analyses at the level of a statement that merely suggests victimisation occurred because the capable guardian was absent. In fact, Pratt & Turanovic (2016) present a criticism of this inherent tautology in the routine activity approach, i.e. victimisation is the outcome of the convergence of likely offenders and suitable targets in the absence of capable guardians, thus when these elements co-occur, victimisation is the result. The problem here, as I also account for it in **Study IV**, is that the basic theoretical notions of the routine activity approach, while easy to comprehend initially, ultimately say very little about the abilities or capabilities of the guardian or, for that matter, who could take up the function of the guardian, i.e. who could potentially foil the successful completion of a crime act.

Routine activity approach 'hard-liners' suggest that guardians are to be understood as separate from the target (Hollis, Felson, & Welsh, 2013), even though taking such a position would present clear negative implications for the study of technology-mediated crime. This is the case primarily because interventions into the convergence of offenders and targets by third persons are often difficult to propose given the personal nature of using digital devices (**Study IV**). Given that debates about who (or what) ought to be considered as the 'capable guardian' are ongoing (see e.g. Buil-Gil, Lord, & Barrett, 2021: 289), I discuss this particular element of the routine activities approach further below.

2.2 The role of guardians and guardianship in crime acts

As mentioned above and discussed at length in **Study IV**, the main debate regarding the 'guardian' (or 'guardianship') concept comes down to whether the guardian is always to be considered as separate from the target, e.g. a (random) third person whose presence may deter crime (Felson & Boba, 2010; Hollis, Felson, & Welsh, 2013; Miró-Llinares, 2015), or if there is an argument to be made for protecting one's own person (**Study IV**). Of note here is that Felson & Cohen (1979: 590) mention in a footnote of the original work that "the analytical distinction between target and guardian is not important in those cases where a personal target engages in self-protection". Hence, Felson and Cohen originally never closed the door on the possibility of at least considering self-protection, although later authors have taken this to be a hard rule (see e.g. Miro-Llinares, 2015)[1]. With respect to guardians and guardianship, Felson himself seems to be ambiguous about the clarity of it, supporting or at least not excluding such an interpretation in some sources (Felson & Cohen, 1979; Felson, 2014 (archive interview); Felson, 2019: 619), while underwriting the opposite (Hollis, Felson, & Welsh, 2013) or keeping the account unclear (Felson & Boba, 2010) in others.

In their critical reappraisal of the concept, Hollis-Peel and colleagues (2011) additionally note that the guardianship aspect of the routine activity approach was dormant in terms of a research agenda for a long time. Conversely, the idea of self-protection or, at minimum, the potential for self-protection specifically with respect to cybercrimes was tentatively present from the beginning of such inquiries (Grabosky, 2001; Grabosky & Smith, 2001), appearing alongside discussions of the routine activity approach. In particular, Grabosky & Smith (2001: 29) emphasise that "much computer-related illegality lies beyond the

[1] It must be noted that Felson and Cohen parted ways soon after the publishing of the routine activity approach, where Felson stuck with and further developed the approach and Cohen returned to more conventional approaches to criminology (see Felson, 2008).

capacity of contemporary law enforcement", which means that other institutions as well as a degree of self-help by potential victims plays a role. Other authors have also suggested various forms of guardianship that do not relate the guardian concept solely to third persons, e.g. personal guardianship (Bossler & Holt, 2009; Leukfeldt & Yar, 2016) or guardianship in the form of knowledge and awareness (Hutchings & Hayes, 2009).

Currently, authors tend to mention guardians as separate third persons alongside the self-protective behaviours of a target (see e.g. Buil-Gil, Lord, & Barrett, 2021: 289–290). Thus, guardianship and (self-)protective behaviours tend to be discussed as separates within the context of organisational settings (Buil-Gil, Lord, & Barrett, 2021; see also Stockman, Nedelec, & Mackey, 2016) where applying the concept of "social guardianship" (Spano & Nagy, 2005: 418) is feasible as a consideration. Nevertheless, non-cybercrime studies have shown that guardians need to be available, capable and willing to act (Reynald, 2009) and that only people directly responsible for supervision actually fulfil the role of a guardian (Collie & Greene, 2016). This inevitably brings us back to the notion of 'responsibilisation' and the fact that such responsibilities are, in fact, placed on people within work settings (**Study IV**).

Given that people only spend about 35–50 hours a week working (Eurostat, 2019), most of our time is spent away from work. What is more, we now spend increasing amounts of our time "alone together" (Turkle, 2011), i.e. connected but alone, and home users are considered to be more vulnerable to different cyber risks (Kritzinger & von Solms, 2010). Thus, the arguments for successful intervention by third persons are questionable at best (**Study IV**), inclusive of issues related to the availability of capable guardians against cybercrime in general. In particular, we have to take into account that only 59% of internet users in the EU think they can protect themselves sufficiently against cybercrime (European Commission, 2020). Furthermore, there is an important difference between perceived knowledge and actual knowledge about a topic such as cybersecurity, where greater perceived knowledge or higher perceived response costs can make people less motivated to take actual protective action (De Kimpe, Walrave, Verdegem, & Ponnet, 2021; Bax, McGill, & Hobbs, 2021), not to mention voluntarily doing so for someone else without first receiving a relevant request from them. With respect to home or other private environments, the latter could additionally be considered as subject matter for the privacy-security trade-off debate (see e.g. Solove, 2011), i.e. intruding into someone's private sphere without invitation.

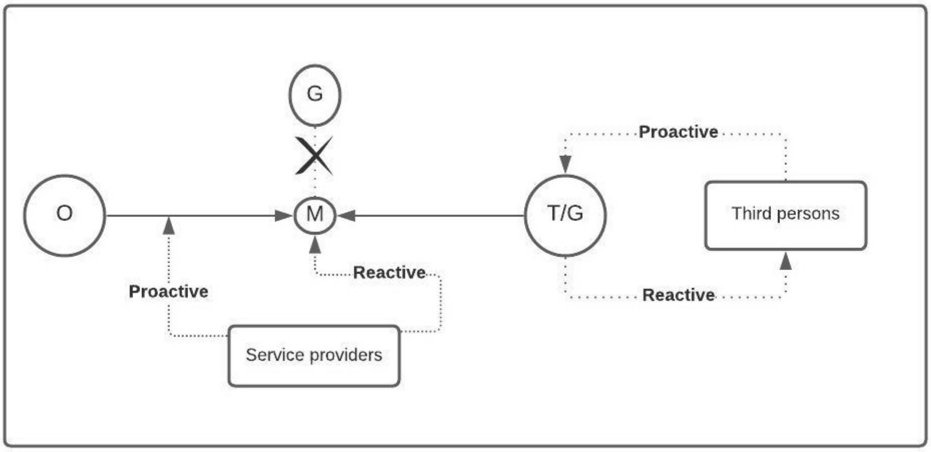

Figure 1. In the absence of third person direct (G)uardians, (T)argets protect themselves from the actions of (O)ffenders in the (M)ediated point of convergence.

Hence, where direct interventions by third persons are excluded from consideration in technology-mediated environments (Figure 1), the outcome of the mediated convergence is determined by an interaction between the offender and its target. The absence of a third person guardian does not, however, automatically mean that crime occurs. For instance, Eck & Madensen (2015) have explicitly suggested considering offenders, targets and guardians as roles that people switch between throughout the day. Thus, where a target and a guardian are considered as roles, it is also possible to propose that a person can carry these two roles simultaneously, i.e. be the target of a criminal offence but also protect him- or herself. Here, it is important to also note proactive and reactive roles available to third persons such as family or colleagues and entities such as service providers. Targets are to be considered "alone" in terms of the immediate contact with the sender. By no means does this suggest, for example, that third persons are not involved in educating the target (proactive) or helping with aftercare (reactive) (Figure 1). Service providers can take technical measures to decrease the overall number of unwanted mediated interactions (proactive) (see e.g. Priezkalns, 2019). Similarly, service providers are able to react to notifications of illegal activities in their networks or services, e.g. block unwanted web traffic where necessary or take down unlawful or harmful content (e.g. Kikerpill, Siibak, & Valli, 2021).

As I mentioned in the introduction, regardless of the verbs we use to describe activities in technology-mediated environments, such activities are always rooted in and dependent upon communication. Thus, it is not by chance that, once third person guardians are removed from the model (Figure 1), the convergence of offenders and target/guardians appears as if depicting a basic communication model. The offender stands for the sender, the target/guardian stands for the recipient, and the mediated point of convergence stands for transmitted messages. **Hence, the crime act in technology-mediated environments is always concurrently a communicative act.**

From the perspective of the offender/sender, the successful completion of the criminal act means that an offender does what he or she intended to do (e.g. levying a threat against someone) or obtains the gains that he or she sought to obtain (e.g. illicit gains from fraud). From the perspective of the target-guardian/recipient, the criminal offence can be prevented in two primary ways, depending on the type of criminal offence. Firstly, where the type of a criminal offence, e.g. a criminal threat or the posting of unlawful content, does not require further action by the target (see e.g. Jõgi, 2012), the only way of preventing victimisation is to close any channels that the perpetrator may use to carry out the crime act. The difficulty of preventing such crimes from occurring is apparent in the modern, ubiquitously connected way of being, including where service providers adopt a passive role (e.g. Drake, 2021) and must often be pressured into action (Aziz, 2020).

The second type of offence, i.e. where further action or compliance is required from the target, can be prevented either by closing any open channels or, if the blocking proves unsuccessful, interpreting the received communication as unwanted, suspicious, criminal etc. and not reacting to the received transmission in a manner expected by the sender. However, criminal actors often manipulate the content and delivery of messages, i.e. socially engineer the communication, to increase the chances of gaining compliance from (unaware) recipients.

2.3 Social engineering then and now

The substantive idea behind the original meaning of social engineering, i.e. large-scale efforts to reshape social structures and reorder society in desirable directions (Fein, 2001: 122), is undoubtedly old. For instance, various reforms and legislative efforts in Ancient Rome have been discussed from the perspective of social engineering (see Sirks, 2013). While the practices are as old as human society, the terminology is newer. As a specific term, "social engineering" emerged from an analogy in which mechanical engineering was compared to bringing about social change in the mid-19th century (Gray, 1842: 117; Hatfield, 2018), i.e. how politicians could be likened to mechanical engineers trying to fix a steam engine. As Alexander & Schmidt (1996: 1) note, "the word engineering suggests the designing and erecting of structures and processes in which human beings serve as raw material".

According to Patel & Reichardt (2016: 9), the rise of social engineering in its original usage coincided with the "incremental creation of the modern interventionist and welfare state". Following its emergence, the concept was further developed and understood through the prism of social and political change (e.g. Addams, 1914) and enjoyed large-scale optimism in the form of "targeted change" in the first half of the 20th century (Patel & Reichardt, 2016). According to Hatfield (2018: 104), social engineering terminology formed a part of various discussions throughout most of the 20th century, but waned in popularity sometime in the end of the Cold War. While the use of specific terminology may have waned, it is important to note here that what Garland described as a "responsibilisation strategy" (see above; Garland, 1996) can be considered as falling under the category of social engineering. Thus, by proxy, the adoption of the NIS (network and information security) directive (Directive (EU) 2016/1148), and large-scale legislative action in general, are instances of social engineering in the original meaning of the concept, because such legislative action aims to "reorder society in desirable directions" (Fein, 2001: 22). Understandably, the notion of "desirable directions" suggests that such chosen directions may be desirable for some but not others.

As Hatfield notes, in the mid-1970s, the phrase "social engineering" was also introduced into the emerging 'hacker culture' (Draper, 2001 as cited in Hatfield, 2018: 105). However, the usage of the concept of social engineering in early hacker communities was disconnected from the concept's original meaning within political science (see Hatfield, 2018: 105). Put differently, the definition of social engineering as used in current cybersecurity literature (see above, Hadnagy, 2018) did not emerge from the original, political science approach to the concept. Nevertheless, the historical/political science understanding of social engineering (Fein, 2001) and the modern cybersecurity understanding of social engineering (e.g. Hadnagy, 2018) share some common traits (Hatfield, 2018). Hatfield suggests that these common traits include:

- epistemic asymmetry, i.e. possessing superior knowledge;
- technocratic dominance, i.e. possessing superior technical knowledge, which allows to enact changes "in the behaviour of others" (Hatfield, 2018: 104);
- teleological replacement, i.e. replacing the purpose of the target of social engineering with one's own.

Although Hatfield's (2018) arguments on the connections between the concept of social engineering as understood either in political science or cybersecurity provide valuable insight, he seems to also avoid addressing the darker side of the concept's original manifestation. This missed notion is, for example, clearly present in the work of Patel & Reichardt (2016) when addressing Nazi social engineering in the 1930–1940s. Furthermore, Conroy (2017: 15) notes that defining or implementing the concept of social engineering was difficult until the rise of Nazism, which empowered "a political eugenicist more radical and extreme than even the most enthusiastic pioneer of social engineering could have imagined"[2]. Hence, the history of social engineering shows that the

[2] I note that the eugenics movement was historically influenced by Renaissance, Enlightenment and 19th century criminological thought, in particular by physiognomy (G. della Porta in the 16th century and J. K. Lavater in the 18th century), phrenology (F. J. Gall and J. Spurzheim in the 18th and 19th century) and the broad idea of "the criminal man" (C. Lombroso) (see Posick, 2018). The social engineering of the mid-19th to mid-20th century eugenicists, most notably in Nazi Germany (see Conroy, 2017), is not merely a horrific mistake of history – the forced sterilization of women continues today as a stark reminder of what came before and has not changed (see e.g. Patel, 2017).

concept has always had both the potential for progressive use (Hatfield, 2018) and horrific abuse (Conroy, 2017). Therefore, the manifestation of social engineering within current social psychology literature more broadly (e.g. Cialdini, 2009), and in the ‘human hacking’/cybersecurity literature more specifically (see e.g. Hadnagy, 2018), exhibits the same traits. Not all social engineering carries an entirely negative connotation (Hadnagy, 2018), because the use of influencing techniques are as much present in parent-child relationships as they are in cybercrime. In fact, the parental use of instrumental lying (Heyman, Luu, & Lee, 2009) is all too common as well as a very familiar phenomenon to most people[3].

Furthermore, I contend that the notion of teleological replacement (Hatfield, 2018: 104) ought to be recast as influencing persons’ will to act, because where influencing fails, no purpose replacement can be said to have occurred. This is supported by Alexander & Schmidt (1996: 2), who note with respect to the original meaning of social engineering that “social engineers have often resorted to violent measures to break the will of those whom they wanted to change”. Moreover, recasting teleological replacement as influencing persons’ will to act gives us (through the notion of influencing) a direct path to analysing how the influencing occurs and what it comprises. For instance, within the public and political sphere, we find discussions regarding the dissemination of propaganda or the use of demagoguery (Roberts-Miller, 2019) as well as “nudging”, i.e. the use of behavioural, economic and psychological insights to influence the behaviour of policy targets (Moseley, 2020; Schmidt & Engelen, 2020). Within the private or individual sphere, we find discussions about the specific uses of influencing techniques (Cialdini, 2009) that are employed to direct people’s behaviour. As my thesis follows the line of current cybersecurity discussions of social engineering, the aforementioned difference in the scale of influencing is crucial as it also reveals the underlying goals of social engineering practices.

Hence, the political science treatment of social engineering primarily concerns itself with (perceived) social problems addressed on a larger scale, e.g. through public policy (Fein, 2001). The goal of social engineering in the cybersecurity domain, however, is mainly focussed on how the behaviour of individuals is manipulated through influencing techniques (Hadnagy, 2018). Admittedly, easy access to ICTs has allowed the individual instances of social engineering to grow into a large-scale problem (PurpleSec, 2021). Even so, this larger problem must be understood as the sum total of a significant number of isolated individual manifestations of social influencing used for illicit purposes.

Thus, it is feasible to begin the study of the larger problem of social engineering attacks by understanding individual instances of it. In particular, important insight can be provided by understanding how (successful) influencing is achieved within individual instances of social engineering, i.e. how

[3] For in-depth accounts of the role of instrumental lying as it marks the concept of ‘social hypocrisy’, see Denery (2015) and Douglass (2020).

the use of influencing techniques appears in social engineering attacks. This approach is also supported by the universal nature of compliance and the fact that no one demographic is more or less vulnerable to the deceptive practices of social engineering (Norris, Brookes, & Dowell, 2019). Given that influencing occurs through communication, a reasonable point of departure in the analysis of crime-as-communication is to focus on the point of convergence that connects perpetrators (of any background and intention) to their targets (of any demographic or disposition) – the messages that perpetrators use to try and change the behaviour of their targets. Furthermore, assessing techniques of social influencing (Cialdini, 2009) as these appear in socially engineered messages is also independent of the size and type of the particular audience or the background of the senders.

Hence, understanding technology-mediated crime as concurrently constituting a communicative act means that the successful completion of the crime act from the perpetrator's perspective depends on the efficacy of influencing techniques employed in the socially engineered messages. In contrast, and from the recipient's perspective, the ability of detecting instances of social engineering within technology-mediated acts of crime(-as-communication) enables the mitigation of risks and the possible prevention of victimisation.

2.4 Towards the crime-as-communication approach

As mentioned in Chapter 2.1, the routine activity approach suggests that crime is likely to occur when a likely offender converges, in space and time, with its target in the absence of a capable guardian. This basic separation of analytical elements of a crime act is broad enough to provide a starting point for the study of various types of crime, including technology-mediated crimes. However, as I noted in detail in Chapter 2.2 and in **Study IV**, there are significant problems with the guardian element, in particular with respect to technology-mediated criminal activity. More specifically, the manner in which people use modern ICT devices suggests that if the guardian is to be understood as a third person, then guardians are almost always absent from the (attempted) crime act. Nevertheless, where a guardian is to be understood as a role (see Eck & Madensen, 2015), then a person being targeted in a crime act can concurrently perform as their own guardian.

Understanding that the individual being targeted for a crime can perform the role of a guardian (see above, Figure 1), and acknowledging that all mediated activity, by definition, is rooted in and dependent on communication, leads us to see the perpetrators as senders, targets as recipients and the point of convergence as the messages exchanged. In other words, the previously described synthesis makes it possible to translate and place criminal activity, which occurs in mediated environments, into a communication-based framework, i.e. the **crime-as-communication** approach.

At this juncture, it is important to note certain aspects that form the crime-as-communication approach:

- there are criminal acts that do not require interaction from the target for successful completion, e.g. robberies (unlawfully taking an object with the use of violence). With respect to mediated crimes, we could therefore refer to "robbery-type" crimes in which the crime target's involvement has no significant bearing on the outcome of the crime act (see **Study I**). Non-technical robbery-type crimes can include, for instance, giving false testimony (see Jõgi, 2012), since the act of lying under specific circumstances is sufficient for committing the offence. Technical robbery-type crimes can include, for instance, cyber-attacks in which transmissions are targeted at machines rather than humans (see e.g. McGuire & Dowling, 2013: 5);
- there are criminal acts that require interaction from the target, e.g. fraud (causing proprietary damage to another person by knowingly causing a misconception of existing facts) or extortion (coercion of another person to transfer proprietary benefits, among other things, through threatening to restrict the liberty of a person). With respect to mediated crimes, we could therefore refer to "fraud-type" or "extortion-type" crimes in which the crime target's response to the perpetrator is influenced either through deception or coercion (**Study I**);
- obviously, not all communication pertains to criminal activities. The dissemination of harmful information (see Hansson et al., 2021) comes in many forms, e.g. disinformation. While not always criminal under applicable law, the spread of harmful information can, and often does, create improved conditions for the success of crime-as-communication. Hence, these other instances of harmful or misleading information can act as context for actual criminal activity.

With respect to crimes that require interaction from the target/recipient, social engineering (see 2.3) is used in an attempt to shape the recipient's response in a way that benefits the sender. This means that the outcome of these types of mediated crime acts is determined by, on one hand, how a perpetrator produces a meaningful message through encoding (Hall, 1973) and, on the other hand, how the recipient meaningfully decodes the message. This is the basic premise of crime-as-communication, i.e. criminals communicating with their targets in a meaningful way. For instance, it would be difficult if not impossible to communicate a lottery win to someone, who has no concept of what lotteries are.

Furthermore, the particular meaning of exchanged messages relies on the context in which said messages were exchanged (Rigotti & Rocci, 2006). In addition to its interpretive dimension, i.e. helping us understand what was communicated to us, context also has a constitutive dimension (Rigotti & Rocci, 2006). The constitutive dimension of context means that, given a specific context, certain messages are chosen over others for the creation of meaningful exchange. Referring back to opportunity theories of crime (see 2.1), context is to messages what immediate circumstances are to an act of ('terrestrial') crime. Put differently, similar to how immediate circumstances (e.g. angry, intoxicated and in a verbal dispute) can contribute to the realisation of a criminal act, the context in which messages are received can contribute to the particular way in which said messages are understood by the recipient (see **Study II, III**). For instance, a text message that notifies a person of a fine that was imposed on them due to a violation of lockdown rules (Salisbury, 2020) can have a particular meaning within the context of the COVID-19 pandemic, but be meaningless otherwise (cf **Study II, III**).

Nevertheless, meaningful exchanges and feasible contexts are ways of ensuring that the transmitted content is understood (as close to) as intended (as possible). People, by default, assume that communication is truthful, because this disposition facilitates interaction in a social world (Levine, 2019). Provided that intentional deception is the exception not the rule (Levine, 2019), people are also more vulnerable to instances where deceptive practices, such as social engineering, are in fact employed. In order for suspicion to arise about certain acts of communication, there must be a triggering event, i.e. something in the message must constitute 'diagnostically useful information' that triggers a person to suspect deception (and, consequently, the potentially criminal nature of the communication). Hence, using socially engineered messages in (more or less) feasible contexts constitutes the perpetrators' toolbox in crimes-as-communication where interaction is required from the recipient. In contrast, the knowledge and ability to detect deception in socially engineered messages is the key to avoid providing the perpetrators a response they seek, e.g. personal information, bank information, access to the device etc. Thus, merely understanding the manifest meaning of the content in communicated messages is not enough to prevent victimisation – some element, i.e. 'diagnostically useful information', within the message and/or its context must trigger the recipient to suspect deception.

2.4.1 Research questions

Following from the aim of my thesis and the theoretical framework provided above, my thesis is guided by the following research questions (RQ):

1. What diagnostically useful information is available to a recipient from the content and context of social engineering attacks? (**Study I, II, III, IV**)
    a. How can conventional crimes defined under criminal law inform the communicative aspects of criminal activity in mediated form? (**Study I**)
    b. Which general communicative approaches appear in the production of messages used in social engineering attacks? (**Study I, II, III**)
    c. What role does impersonation play in social engineering attacks? (**Study I, II, III**)
    d. Which media are used by criminals to perpetrate social engineering attacks (**Study I, III, IV**)?
    e. What topics and themes do criminals rely on in social engineering attacks? (**Study I, II, III)**

    f. Which specific social-psychological influencing techniques are employed in social engineering attacks? (**Study III**)
    g. What role does social context play in social engineering attacks? (**Study I, II, III**)
2. How does the crime-as-communication approach contribute to current criminological thought on cybercrime? (Entire thesis)

3. METHODOLOGY

In this Chapter, I provide an overview of the methods that guided my research in the three empirical studies (**Study I, II, III**), which motivated the theoretical criticism put forward in **Study IV** and, ultimately, lead to the development of the ***mazephishing*** framework (**Study III**) and the concept of **crime-as-communication**. The overview of the main data collection and data analysis methods used in the empirical studies is provided in Table 1.

Table 1: Overview of the data collection and data analysis methods used in the thesis.

	Data collection	Data analysis
Study I	Emails (n=297) received in two personal email accounts	Qualitative text analysis (Ezzy, 2002) with coding scheme developed on the basis of applicable criminal law provisions
Study II	News stories (n=831) from international media	Qualitative content analysis (Krippendorff, 2004)
Study III	News stories (n=563) from international media	Qualitative and quantitative content analysis (Krippendorff, 2004)

Following from the focus on messages as points of mediated converge between senders (perpetrators) and recipients (targets), I used qualitative text analysis in **Study I** and content analysis in **Study II** and **III**. Although the general approach is consistent and similar across all three empirical studies, each work built on the previous one. As a researcher, I am primarily interested in the storytelling and persuasion abilities of scammers as these abilities are reflected in the messages that ultimately reach recipients. Since there are numerous aspects that researchers can focus on within the broader field of susceptibility to social engineering attacks (see Nguyen, Jensen, Durcikova, & Wright, 2020: Appendix A), my background in law was a determining factor at the start of **Study I**.

Doctrinal legal research has developed its own rules for interpreting social phenomena and colleting "data", i.e. binding normative sources, principles of law, case law and authoritative scholarship (see Van Hoecke, 2011). Due to this significant difference in its approach to research data, doctrinal legal research is essentially bound to produce new opinions instead of new facts. Nevertheless, the importance of the *nullum crimen sine lege* principle (no crime without applicable law) cannot be entirely set aside when departing on a study of crime. This is why **Study I** uses data suitable for a systematic analysis of a social phenomenon but applies a coding scheme developed on the basis of applicable Estonian criminal law, i.e. takes a socio-legal approach (Van Hoecke, 2011). Hence, it is possible to trace how my thinking and approach to research developed from **Study I** to **Study II** where I set legal aspects aside in lieu of a move towards the disciplines of criminology and communication. In particular, this move brought with it a more generic approach to the study of crime as a socially constructed phenomenon (Posick, 2018). Moreover, setting aside specific legal aspects and the strict frame of legal interpretation also provided more creative freedom in both data collection and its subsequent analysis. This opened the opportunity to focus on the content as well as the context of social engineering attacks and develop the *mazephishing* framework (**Study III**). Thus, the crime-as-communication approach is the result of the previously described step-by-step developments from a socio-legal research (**Study I**) to qualitative content analysis (**Study II**) to a hybrid (qualitative and quantitative) content analysis (**Study III**).

Below, I provide a more detailed overview of the substantive choices regarding data collection and data analysis methods employed in **Study I**, **II** and **III** using a study-by-study structure to better reflect the previously described developments.

3.1 Study I

The aim of **Study I** was to ascertain the prevalence of e-mail based social engineering attacks perpetrated against an individual (in this case, the author of this thesis) over a one-month period from mid-August to mid-September in 2018. Emails used as data (N = 297) were collected from two email accounts of the author and the final sample used in the subsequent analysis comprised emails (N=42) that were in English and exhibited strong initial indicators of a phishing attack. These initial indicators include, for instance, an unknown sender, grammatical errors, subject lines with upper-case letters throughout as well as ambiguous, generic or overtly out of place topics (Jakobsson, 2007: 3–6).

Due to its exploratory and socio-legal nature, **Study I** used qualitative text analysis with open coding as the method with one e-mail as the unit of analysis (Ezzy, 2002). The chosen qualitative approach allowed me to explore the relatively small final sample of emails in more depth. The guiding concept of coding in **Study I** was the influence and impact that message content could have on recipients' will to act, which accounts for the fact that the ultimate goal of social engineering attacks is to influence the recipient into action (Khonji, Iraqi, & Jones, 2013).

I developed the coding scheme on the basis of existing phishing literature, e.g. the use of 'urgency cues' in social engineering attacks (Williams, Hinds, & Joinson, 2018), and concepts used in the legal analysis of certain criminal offences against property (robberies, extortion and fraud) (see RQ 1a). I made use of how legal analysis approaches the influencing of a person's will to act (Sootak, 2010) for:

- robberies – a person's will is broken, i.e. compliance is not expected or required from the target;
- extortion – a person's will is bent through coercion; and
- fraud – a person's will is influenced through the creation of misconceptions, i.e. a person makes the decision, but based on misleading information.

On one hand, robberies and extortions exhibit a more intense influence on a person's will to act, while this intensity is lesser in cases of fraud. On the other hand, robberies do not expect or require interaction from the target for the completion of the crime act, while extortion and fraud do. For this reason, I completed the four-element scheme by adding an informational-type interaction (**Study I**: 51), which is considered low in intensity and does not expect or require specific compliance from the target.

The notion of specific compliance relates to an understanding of control within the interaction that results or can be expected to result from the social engineering attack. For instance, the idea of robbery-type communications was based on the conventional criminal offence of robbery in which the perpetrator overpowers the victim entirely, e.g. knocks a victim unconscious before taking their belongings. A hypothetical example from the digital sphere would be an instance where, upon opening an e-mail merely for reading, the target's device is overtaken or locked without any further interaction expected or required from the victim, e.g. a ransomware attack (Proofpoint, 2019) that does not require the recipient to open any attachments.

Similarly, while there may be implicit notions of compliance embedded in the message in informational-type instances, the sender lacks control required for specific compliance, i.e. getting the recipient to visit a specific website or open a specific e-mail attachment. Thus, if the sender lacks a way for eliciting specific compliance, the perpetrator also lacks control over the attack's outcome.

Thus, the expectation for **Study I** was that the final sample (N=42) would present some type of split between extortion- and fraud-type communications. While extortion-type and fraud-type differ in terms of intensity on a person's will to act, i.e. use of threats *vs* use of deception, both require 'specific compliance' from the target of the message.

Finally, it is important to note that although the sample in **Study I** was limited for specific reasons, i.e. data collection only covered a particular time period and the sample was not supplemented with emails from sources other than the email accounts (compare with Atkins & Huang, 2013), this does not negate or exclude the possibility of scaling the research design for representative samples in the future. However, since legal interpretation, i.e. determining the type of offence committed or attempted, is a manual process that requires specific knowledge, the manpower required for scaling the design used in **Study I** would likely be extensive. Even so, using a similar basic research design in the future would be feasible for gaining insights into the cybercriminal activity faced by smaller (or highly particular) populations. Furthermore, basing the coding design on actual legal provisions from applicable criminal law also allows researchers to compare their study results with available official crime statistics, which is not possible with a generic treatment of crime acts.

3.2 Study II

The starting point for **Study II** was a gap in existing literature, which suggested that while the use of salient current events as the interpretive backdrop for social engineering attacks is generally recognised in publications and reports of the cybersecurity industry (RiskIQ, 2020), the same cannot be asserted for academic scholarship (Verma, Crane, & Gnawalli, 2018). After noticing an increase in reports of COVID-19-themed social engineering attacks in well-known specialist newsletters such as The Cyberwire and The Hacker News in early 2020, the aim of **Study II** was to understand the extent to which fraud- and extortion-type scam messages appear "in the wild", i.e. in non-controlled conditions, within and as inspired by the context of the COVID-19 pandemic (see RQ 1b & 1g).

The data collection for **Study II** relied on publicly available information and started with an initial sample (N=1928) of results from a Google keyword search using the phrases "covid scam", "covid phishing", "coronavirus scam" and "coronavirus phishing", while limiting the results to ones appearing in English. Making use of Google's search operators, these four phrase-based searches were performed separately for each month from January 2020 to April 2020. To further facilitate the data collection and initial analysis, I used the Linkclump Google Chrome extension, which allows to collect the links of search results instead of immediately going through each result.

Using qualitative content analysis with some elements of quantification (Krippendorff, 2004), the initial data were evaluated and separated into categories of "mainstream media" (e.g. online newspapers and websites of different radio and TV news stations) and "specialist media" (e.g. online news platforms dedicated to technology and cybercrime reports such as BleepingComputer or Tech Xplore).

Two rounds of exclusions were made from the initial sample. Results that only contained video, appeared as posts on social media or were published on websites of private companies (e.g. banks or law firms), on the official websites of public authorities and different non-profit organisations were excluded from the final sample. These initial exclusions were made to ensure consistency in data sources, i.e. text-based news stories. Further, by relying on the communication types developed in **Study I** (robbery-, informational-, fraud- and extortion-types), news stories that pertained to hoaxes and misinformation (informational-type) or purely technical attacks in which the exploitation of human vulnerabilities would be a non-issue (robbery-type) were also excluded (see 4.4.).

Following exclusions, the (thematic) content analysis was carried out based on the final sample of news stories (*N*=831) as appearing in mainstream media (*N*=618) and specialist media (*N*=213). To capture the variety of communication strategies present in the social engineering attacks in the sample, my co-author and I used single words, phrases as well as sentences and paragraphs as units of coding. Most importantly, we discussed the meanings of codes developed on the basis of **Study I** and shared our interpretations throughout the process. The primary guidance related to the gain-based nature of fraud-type communications, i.e. scam messages the content of which attempts to deceive recipients into action by offering to fulfil certain of their needs, and the loss-based nature of extortion-type communications, i.e. scam messages that use fear appeals and threats of loss to coerce recipients into action. These communicative strategies were tentatively called "the Good Samaritan" (based on the fraud-type) and "Shock and Awe" (based on the extortion-type).

The Good Samaritan strategy can be expected to see use where perpetrators aim to respond to a (sudden) demand, including material or psychological desires, of a target with a fictional supply. Thus, the Good Samaritan strategy is to be understood as a reward or gain-based approach, i.e. promising the target of the social engineering attack something that the target is already seeking or may be interested in obtaining. Such needs may include companionship (e.g. Carter, 2021), money (Carter, 2015) or certain goods, e.g. personal protective equipment during the COVID-19 pandemic.

The Shock and Awe strategy, conversely, emerges from using fear appeals and urgency cues to suggest to a target that they may experience some sort of loss, e.g. financial or reputational, unless action is taken. Unlike the Good Samaritan strategy, which aims to lull the target into an expectation of gain, the Shock and Awe approach confronts the target directly with the possibility of a penalty or loss (see Williams & Polage, 2019), e.g. by threatening to release sensitive materials or shut off electricity unless payment is made.

3.3 Study III

The data sample used in **Study II** was also employed in **Study III**. However, certain important coding specifications were developed and added, and a further layer of quantitative analysis was implemented in **Study III** as described below. For **Study III**, I made further exclusions from the sample of news stories established in **Study II** (N=831), in particular with respect to excessive repetitions of essentially identical news stories. Hence, the final sample of news stories for subsequent content analysis in **Study III** was N=563. Furthermore, since some news stories covered more than one scam, each scam within a news story was counted separately, resulting in N=1040 scams.

Where **Study II** had a more qualitative emphasis with minimal elements of quantification to ascertain the overall presence of the "Good Samaritan" and "Shock and Awe" approaches, important specifying coding elements were added in **Study III**. In addition to coding the employed approach to message production, the coding scheme used in **Study III** included:

- the person of the sender of a scam, including whether international organisations, public authorities or private companies were impersonated (RQ 1c);
- the theme used in the scam, e.g. health information, provision of goods or soliciting donations (RQ 1e); and

- the medium used for delivering the social engineering attack, i.e. emails, text messages, phone calls, social media messages or posts and websites (RQ 1d).

Thus, **Study III** used a hybrid approach to content analysis, i.e. combining elements of qualitative and quantitative content analysis. Furthermore, the analysis presented in **Study III** made use of Cialdini's (2009) six principles of persuasion to qualitatively assess the content of previously coded descriptions of social engineering attacks (see RQ 1f). These six principles of persuasion proposed by Cialdini (2009) are:

- authority – people exhibit a tendency to comply with requests from authoritative figures;
- scarcity – people desire an item more when there is competition for it or its availability decreases;
- liking and similarity – people are more easily persuaded by someone they know and like;
- social proof – people look to others to confirm their decisions;
- reciprocity – people are inclined to return a favour provided to them;
- consistency – people aim to be consistent in their actions and keep to their commitments;

From the combination of the coding scheme, parts of which were developed for **Study II** but significantly specified and added to in **Study III**, and the use of Cialdini's six principles of persuasion, I developed the messages-in-context analytical framework of *mazephishing*. The *mazephishing* framework looks at three elements that characterise a social engineering attack: the context of the social engineering attack, e.g. salient social circumstances that have or may have an impact on the content of scams, the medium that is used for delivering the attack (i.e. the message), e.g. emails, text messages or phone calls, as well as the influencing techniques that appear from the content of the scam.

3.4 Reflecting on the advantages and limitations of the chosen methodology

Focussing on the content of messages provides certain privileges over other methodological choices, but also presents some limitations. Firstly, analysing messages rather than interactions or message exchanges between senders and recipients takes away the possibility of investigating the development of the perpetrator-target relationship over time (see e.g. Carter, 2021). Nevertheless, Carter (2015) showed that even scam leaflets delivered by conventional mail have the potential to deceive recipients into action. In essence, even a single message and the intricacies of its content are sufficient for gaining insight into the strategies used by (cyber)criminals.

Secondly, while focussing on message content and context does not provide accounts of the actual reception of such messages, i.e. the accounts are limited to the analysis carried out by a researcher, it gives the opportunity to take an in-depth look into the ways messages differ, for example by topic/theme, context, medium used etc. This opportunity is in direct contrast with the limitations of most susceptibility studies (see e.g. Nguyen, Jensen, Durcikova, & Wright, 2020: Appendix A), because checking for susceptibility mostly means using a relatively small sample of crafted messages to ascertain recipient reactions.

Thirdly, and specifically in contrast with the data collection choice made in **Study II** and **III**, the primary focus in social engineering attacks' research is on phishing emails (Nguyen, Jensen, Durcikova, & Wright, 2020), as these are the most plentily available form of data. According to Button & Cross (2017), while people are often aware of some types of scams, they can concurrently be completely unaware of other types, including the media used for scams. Thus, solely focussing on emails (e.g. **Study I**) only solves part of the puzzle with respect to understanding the approaches used in socially engineered messages. Due to their availability, emails are a good starting point for larger-scale investigations into message-strategy developments over time (Stojnic, Vatsalan, & Arachchilage, 2021), but more information is needed about other media used in scams such as phone calls (see e.g. Armstrong, Jones, & Namin, 2021). Hence, although **Study II** and **III** focus on content and not interactions or reception, these works also cover numerous media, which provides an improved starting point for future studies into the content of social engineering attacks, including as influenced by salient social contexts.

Finally, the preference of qualitative inquiries over quantitative approaches, i.e. with the exception of **Study III**, stems from two primary reasons of which the first relates to the discipline of criminology in general and the other to the overall debate regarding the method of content analysis. The study of cybercrime is often avoided by criminologists, in some cases due to perceived technical complexity (Diamond & Bachmann, 2015; Maimon & Louderback, 2019). Furthermore, as quantitative approaches loom large over qualitative inquiries (Jacques, 2014), the knowledge that is being created lacks intensity for properly understanding the nuances of what goes on at the core of cybercriminal convergence, i.e. in the messages used to commit these crimes.

This also relates directly to the second concern, i.e. the Berelson-Kracauer debate (Schreier, 2014: 171) in which the “manifest content of communication” approach (Berelson, 1952: 18, as cited in Schreier, 2014) was contested by Kracauer (1952, as cited in Schreier, 2014: 171), who argued that meaning is often complex, holistic and context dependent, and “is not necessarily apparent at first sight”. Hence, while quantification certainly played an important role in **Study III**, the utility of the approach was enabled by the qualitative basis already established in **Study I** and **II**.

4. FINDINGS

4.1. Sender identities

Since every activity, including criminal activity, in technology-mediated (online) environments is rooted in and dependent upon communication, the convergence between a sender (the supposed criminal actor) and the recipient (the potential victim) manifests as an exchange of messages. In modern technology-enabled communications, such as e-mails and text messages, the recipients' attention is first drawn to some identifying information about the sender, e.g. their name. However, it is relatively easy to "spoof" the sender's identity in technology-mediated communications (**Study I**), i.e. make it seem like the incoming communication is from someone that has actually nothing to do with the message (see RQ 1c).

The emails analysed in **Study I** showed that it is common for perpetrators of social engineering attacks to use fake names, including internationally well-known company names such as JP Morgan Chase (**Study I**). Thus, as a common form of deception, "spoofing" allows senders to take on (perceived) identities that may mean something or seem authoritative to recipients. For instance, in the 1040 scams analysed in **Study III**, a total of 1019 (98%) used faked sender information. Due to its easy application, and the fact that people are inclined to comply with requests from authorities (Cialdini, 2009), impersonating specific entities or persons allows perpetrators the opportunity to try and deceive recipients with the added benefit of appearing as a source of authority. **Study I** showed that such fake sources can include banks and government entities, while **Study II** and **III** further added international organisations (the World Health Organisation), healthcare institutions (Center for Disease Control), tax authorities (Her Majesty's Revenue and Customs), local utilities companies and charitable organisations.

Although **Study I**, **II** and **III** showed that perpetrators often use well-known fake identities for added perceived credibility, it is not always a necessary element of socially engineered messages. For example, the criminals also use entirely made-up names and even leave the sender's identity anonymous (**Study I**: 56–58). The name "Iris J. Stobbs" was provided in a scam suggesting the availability of millions of inherited dollars and the sender was "anonymous" in an email threatening to release sensitive materials featuring the recipient. Thus, senders' so-called choice of identity seems to follow from the topic and content of the message: well-known "spoofed" identity for messages relating to some potentially relevant aspect in a recipient's daily life (banks, utilities companies, healthcare institutions etc.), made-up names for completely unexpected communications (**Study I**: 55), or even seemingly anonymous senders where this ambiguity serves a purpose, e.g. creates an aura of mystery around the criminal's identity.

4.2 Hot topics and social context

Beyond the identity of senders, the recipient can also determine the general topic of the message. For instance, this may be available from the subject line of an email (**Study I**), but primarily comes from the content of the message itself (**Study I**, **II**, **III**), because not every medium provides the opportunity of giving a “title” to the message, e.g. phone calls.

The research design of **Study I** meant that messages forming the sample were collected “as is” (**Study I**: 50), i.e. no topic related categorisation or filtering was applied to data collection. This provided a sample of emails anyone could receive in the course of their normal daily lives, which was important for the purposes of capturing the presence or absence of different kinds of message topics. Therefore, the messages were from “rich widows” (**Study I**: 57–58), who have fallen ill and want to give away their inherited wealth, as well as so-called expert “hackers” attempting to extort the recipient by threatening to release sensitive materials. The first of these is a typical example of an advance-fee fraud where a recipient is offered a significant amount of money in return for a small (monetary) favour, e.g. paying the transfer fee of the larger sum[4]. The second example is that of “sextortion” (sex and extortion), which is used to coerce recipients into paying what the sender demands (**Study I**: 56–57).

These generic scam topics often appear in waves, i.e. as fraudulent campaigns where identical or substantively similar messages are sent out in short succession. For instance, the sample of **Study I** included three almost identical sextortion messages (**Study I**: 56). As with their chosen “identities”, scammers are also free to choose the topics around which scam messages are constructed. While these topics are often generic for the purposes of capturing as wide an audience as possible under normal social circumstances (**Study I**), certain events or phenomena in society prompt perpetrators to choose particular topics and themes (**Study II**, **III**) (see RQ 1e).

In **Study II** and **III**, I therefore focussed entirely on scams that were spurred on by the social salience of the COVID-19 pandemic. To study the way in which important social circumstances, i.e. social context, impacts the topics and content of scam messages (see RQ 1g), I created the *mazephishing* framework. The name *mazephishing* was inspired by an old fishing technique called ‘almadraba’ (Arabic for ‘a place to smite’), which involves setting up elaborate underwater mazes of nets to catch tunas during their migratory journeys through the Strait of Gibraltar. Thus, the catch depends on (1) proper timing, i.e. knowing why fish are on the move in large quantities at certain times, (2) place, i.e. interrupting the tunas’ movement at a location and in a manner suitable for the fishermen, and (3) trap-setting technique (**Study III**). Hence, the first element

[4] These scams are also referred to as “Nigerian 419 scams” or 419 scams and the reference to 419 comes from the fact that section 419 of the Nigerian Criminal Code is for the offence of fraud (Chiluwa, 2019). Brody, Kern, & Ogunade (2020) provide an in-depth history of the development of 419 scams.

of the *mazephishing* framework analyses the social timing of technology-mediated crime.

Study II and **III** showed that the COVID-19 health crisis provided scammers with ample ideas on how to exploit people, who were forced to switch to remote work and relied heavily on digital communications. For instance, in the early months of the pandemic, scammers preferred sending out fraudulent messages claiming to provide important health information, which seemingly came from local welfare agencies (**Study II**), the World Health Organisation or the Center for Disease Control (**Study III**). As the social circumstances developed, other topics such as financial relief (e.g. government aid or supermarket coupons), offers for difficult-to-obtain personal protective equipment (e.g. masks) and even cures and vaccines were added as topics (**Study II**, **III**). Based on the themes and topics employed in scams across the first four months of the COVID-19 pandemic, **Study III** (see **Study III**: Table 3) clearly showed that criminals stay abreast of actual social developments and adjust scam content accordingly.

Thus, under normal circumstances, e.g. in the absence of a socially important event or phenomenon, the topics used by criminals in social engineering attacks are often generic to capture the attention of the widest audience possible (**Study I**). However, when presented with a relevant opportunity, e.g. a sudden change in social circumstances such as the COVID-19 pandemic, at least some criminals make efforts to adjust the content of the scams they use to deceive and victimise recipients (**Study II**, **III**). Here, simple logic suggests that salient social circumstances offer scammers some advantages in comparison with normal circumstances, because making the effort of adjusting scam content would be meaningless otherwise, i.e. if the expected “catch” would be of a similar size regardless of circumstances.

4.3 General communicative approaches used in scam message production

The aim of social engineering attacks is to guide message recipients towards specific action, e.g. clicking on a link, opening email attachments, making payments to the sender (**Study I**, **III**) as well as providing the sender with personal or financial information (**Study I**). Therefore, any chosen identities or topics are tools for framing the content found in the body of the message.

In **Study I**, for example, a made-up identity was used to suggest that the sender’s health is failing, which made them want to transfer large sums of money into the custody of the recipient for charitable purposes. Given that **Study II** and **III** focussed specifically on scams circulated during the COVID-19 pandemic, comparable examples included providing recipients with free access to Netflix, supermarket coupons and vouchers and early access to financial relief (**Study II**, **III**). The general connection between the aforementioned examples is that the content of the message is created to persuade the recipient into thinking that they will be gaining something from following the sender's instructions. In **Study I**, such messages were categorised as fraud-type communications, since the offence of fraud relies on the creation of a misconception for the purposes of receiving proprietary gain (**Study I**: 51; see RQ 1a). In **Study II**, this idea was further developed into the "Good Samaritan" general communicative strategy, i.e. an approach to scam message production in which the overall deception is based on offering to fulfil some need of the recipient (see RQ 1b).

In contrast to the gain-based approach to scam message production, perpetrators also have the choice of threats and coercion, i.e. taking a fear appeal or loss-based approach in their messages. For instance, this includes threatening to release sensitive materials (**Study I**) or to infect the family members of the recipient with the COVID-19 virus (**Study III**) unless payment is made. In **Study I**, such messages were categorised as extortion-type messages, because the offence of extortion relies on threatening the subject with certain action unless a specific counteraction is performed (**Study I**: 51; see also RQ 1a). In **Study II**, this idea was further developed into the "Shock and Awe" general communicative strategy, i.e. an approach to scam message production in which the overall deception is based on threatening the recipient with the loss of something (e.g. reputation or health) (see RQ 1b). Furthermore, **Study II** also showed that perpetrators threatened recipients with shutting off their electricity connection unless payment was made and also sent bogus fine notifications, which suggested that the recipient had violated COVID-19 pandemic lockdown rules (**Study II**).

Study III showed that 86.5% of the scam messages (N = 900) employed the "Good Samaritan" communicative strategy and 13.5% (N = 140) messages used the "Shock and Awe" communicative strategy (**Study III**: Table 2). Based on these findings, perpetrators preferred to lure recipients with gains rather than threaten them with potential losses during the first four months of the COVID-19 pandemic.

4.4 A variety of media for "setting the traps"

The second element in the *mazephishing* framework (**Study III**) concerns the "place" where the perpetrators attempt to carry out the scams. In technology-mediated environments, the "place" refers to the medium in which the fraudulent communication occurs (**Study III**). Given the easy and (near-)immediate direct access to recipients that technology-mediated communications provide (**Study IV**), people come into contact with criminals often (**Study I**) and without a feasible possibility of intervention from third persons (**Study IV**).

While **Study I** provided an in-depth view on how criminals use the medium of email to circulate different scams, **Study III** used a broader approach by analysing scams (N=1040) circulated via emails (53.5%), text messages (12.6%), phone calls (13.6%), social media posts and messages (3.1%) as well as fraudulent websites (11.3%), including some scams for which a medium was not specified (5.9%) (**Study III**: Table 5; see RQ 1d). This variety of media used by perpetrators to circulate scam messages further stresses the importance of interpreting message content, in particular due to the differing affordances of the aforementioned media. For instance, where emails have headers (i.e. sender information) and subject lines (i.e. potential information about message topic), phone calls or text messages from unknown numbers cannot be assessed in a similar manner. Therefore, the content of messages is a more robust source for so-called diagnostically useful information (Levine, 2019), i.e. information which should trigger the recipient to suspect deception.

Following from the above, **Study III** further established that along with the use of a variety of chosen identities, themes and topics, communication media and general (gain- or loss-based) approaches to message production, the employment of specific social-psychological influencing techniques (Cialdini, 2009) was also detected from the analysed scam messages (see RQ 1f). These so-called "trap-setting" techniques form the third element of the *mazephishing* framework (**Study III**).

Study III provided empirical support with respect to all six principles of persuasion as suggested by Cialdini (2009). For instance, the principle of authority was present in messages where perpetrators impersonated healthcare institutions when contacting recipients under the pretence of providing COVID-19-related health information (see also 4.1). Furthermore, the principle of scarcity was employed in offers of difficult-to-obtain personal protective equipment, but also in offers of non-existent cures and vaccines. The principle of reciprocity was employed in various solicitations of donations for bogus charities to help relieve the financial strain from COVID-19. The principle of liking was employed by a doctor, who was arrested for selling "COVID-19 Treatment Packs", which included hydroxychloroquine, i.e. the anti-malaria drug, which was endorsed by then President Donald Trump but was unproven in the treatment of COVID-19. The same example also made use of the principle of social proof, since Donald Trump was very vocal about his support for the drug and could, therefore, influence his supporters to try the drug for themselves.

The principle of consistency and commitment was clearly present in utilities scams (paying electricity bills on time), fake fine notifications (law abiding citizens admit fault and pay their fines) as well as bogus messages notifying people that they had come into contact with someone infected with COVID-19 (assuming that the notified people find it necessary to provide further information after receiving the message) (**Study III**). Since the latter two examples derive their significance directly from the circumstances of the COVID-19 pandemic, these provided strong support for all three elements of the *mazephishing* framework (i.e. the elements of social context, medium of choice and the presence of influencing techniques meant for creating deception) (**Study III**). Furthermore, the utilities scam related to electricity bills also went through a type of transformation during the COVID-19 pandemic. My analysis of media texts suggests that the scam was changed from a "Shock and Awe" message type that employs the principles of authority and consistency to threaten people with shutting off electricity unless payment is made into a “Good Samaritan” message type that employs the liking principle to offer people discounts on outstanding utilities bills (**Study II**, **III**). It is probable that the aforementioned change was prompted by local regulations that prohibited shutting off essential utilities in emergency situations (WHSV, 2020; **Study III**). This further shows that not only do scammers stay abreast of relevant social circumstances, but also fluently reflect such changes in the scam messages they produce and circulate.

5. DISCUSSION

I began my thesis from the premise that any and all activity in technology-mediated environments is rooted in and dependent upon communication. Some of this activity – certain information exchanges – amounts to criminal activity, i.e. cybercrime.

As I argue in **Study IV**, the nature of technology-mediated communication makes it easy to gain direct and (near-)immediate access to targets, i.e. access without feasible immediate intervention from third persons (**Study IV**), which means that the protective role of third persons is severely diminished in the moment of mediated convergence. Once third persons are effectively (and logically) removed from the convergence of a perpetrator and their target in technology-mediated circumstances, we are left with a crime act that is concurrently a communicative act between the perpetrator and the target. In effect, perpetrators become senders, targets become recipients, and the convergence between the two manifests as an exchange of messages through mediated points of convergence. Senders transmit (push) messages/content to mediated points of convergence, recipients can be 'notified' of incoming communications, and receive (pull) the message/content for interpretation (Figure 2). These basic elements form the core of the **crime-as-communication** approach.

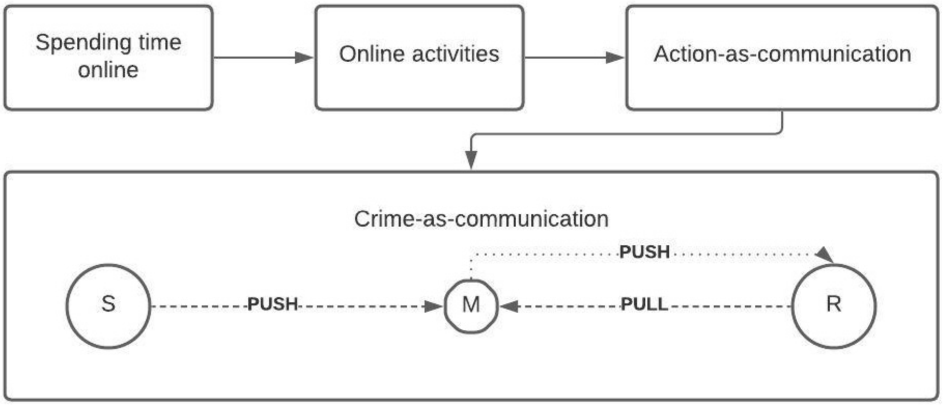

Figure 2. Model illustrating the crime-as-communication approach.

Thus, a significant contribution of the crime-as-communication approach comes from the way the process of crime commission is understood and analysed (see RQ 2). Emphasis on senders, recipients, messages and mediated points of convergence is an approach derived from properly answering the question "What is said when we say *X*?" where X stands for any common description of activities in technology mediated environments, e.g. "using online banking", "using online dating sites", "using social media" etc. However, descriptions of "online activities" and their role in studies of cybercrime victimisation (e.g. Ngo, Piquero, LaPrade, & Duong, 2020) often only amount to what I consider to be "lazy signifiers". Lazy signifiers provide people with a rough idea of what is being signified and can be used for smoother (and more concise) communication on account of shared fields of experience (see Schramm, 1954 as cited in West & Turner, 2019: 11) between sender and recipient, which allow for meaningful exchanges without going into excruciating detail. This approach, however, is insufficient for the analysis of cybercrime and needs more specificity (see Figure 2).

For example, a common finding in current cybercrime-related criminological research is that people are at a greater risk of becoming a victim of cybercrime the more time they spend online (Leukfeldt & Yar, 2016; see also Ngo, Piquero, LaPrade, & Duong, 2020: 433–434). On one hand, this explanation is entirely logical, because if cybercrime is crime perpetrated with the use of connected ICT devices, then actively using such devices is a necessary condition for the possibility of victimisation. On the other hand, this explanation is highly problematic, because a person can "be online" for most of the day on most days of the year (e.g. IT experts or 'hackers') and not become a victim, while a person could also "be online" for 10 minutes only once a month to access their bank account and become a victim of bank fraud. Clearly, this level of explanation provides no real insights about the 'risk' in the supposed risk of becoming a victim. Thus, it is important to notice how little information is conveyed by the phrase "spending time online" if we do not also ask and properly answer the question "What is said when we say 'spending time online'?".

Therefore, the next level of specificity would be to deconstruct "spending time online" into different categories of activity (i.e. "online activities") such as "using online banking", "using social media", "using online dating sites" (see e.g. Ngo, Piquero, LaPrade, & Duong, 2020). However, these signifiers are still too opaque to provide useful insights into why a person becomes a victim of an act of cybercrime.

Hence, a further level of specificity is required, i.e. the action-as-communication level. For example, "using online banking" is, to put it simply, a series of requests for certain information made to, and responses received from, specific servers through the use of one's networked device. In these information exchanges, the user's interpretations of the received information/content form the basis for deciding on the next action (in the interaction). In other words, understanding that "using online banking" is, in fact, a series of communicative acts begins to reveal how this communication process can be initiated, mimicked and exploited by perpetrators for illicit gains, i.e. crime-as-communication is a subcategory of action-as-communication. Common examples include perpetrators impersonating bank personnel to obtain sensitive information (Flinders, 2020) or creating bogus look-alikes of a bank's website to deceive a user into inserting their authentication information (Swinhoe, 2020). Thus, the ways in which people become victims of cybercrime cannot be adequately explained on the level of "spending time online" or "using online banking" and "using online dating sites", but can be explained, for instance, by analysing how people interpret or may interpret incoming communications and the actions such interpretations lead to (Figure 3), including in longer interactions (see e.g. Carter, 2021).

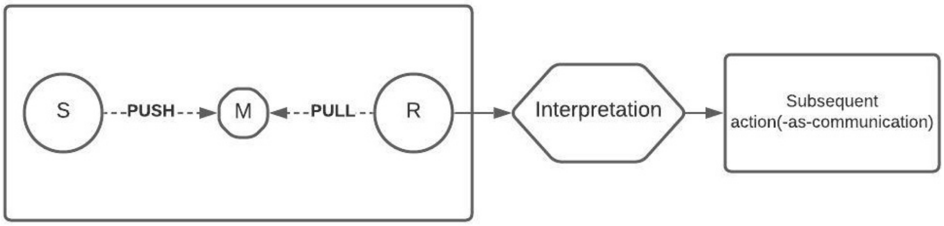

Figure 3. Action-as-communication leading to action(-as-communication).

By making explicit the communication processes that constitute cybercriminal activity, including crime(-as-communication) acts that do not directly target human recipients (see also 4.4 and 5.1), the crime-as-communication approach provides an original contribution to understanding technology-mediated crime specifically, and action in technology-mediated environments more broadly. The latter primarily concerns instances where the use of lazy signifiers, i.e. considering and presenting as a meaningful whole something that is better understood as its constitutive elements, hinders rather than contributes to the thinking and research process.

5.1 Crime-as-communication

As noted above (see also 4.4; RQ 2), not all communication-based criminal activity requires interaction from a human target. For instance, a recipient does not have to further interact with criminal threats in order for the sender's communicative act to also constitute an act of crime. Furthermore, purely technical attacks (see e.g. McGuire & Dowling, 2013: 5) exploit vulnerabilities in machines without the need of involving a human target. In fact, the reason why attacks targeting the human element have increased in importance starting from the 2000s (Jagatic, Johnson, Jakobsson, & Mencer, 2007) is the respective progress in technological barriers to purely technical attacks. Put differently, it is often easier for criminal actors to 'hack humans' (see Hadnagy, 2018) in a social engineering attack than it is to overcome complex technological security measures. Regardless of the target in question – be it machine or human – I argue that it is impossible to deny that cybercriminal activity is rooted in and dependent upon communication.

The above further raises the question whether a distinction between cyber-dependent and cyber-enabled crimes (McGuire & Dowling, 2013), i.e. crimes that are not possible without connected ICT devices (technical attacks) and crimes that are further enabled by connected ICT devices (e.g. fraud), is wholly necessary. For instance, it is possible to ask whether criminal threats levied via phone call, text message or email are, at their core, somehow qualitatively different from threats spoken face-to-face (see also Henry & Powell, 2014; Button & Cross, 2017). It becomes evident that for crimes, which can be wholly or partly perpetrated through communication, relevant communicative aspects must become the primary instances of analysis.

Furthermore, under the crime-as-communication approach, the only actual difference between the previously described offence categories (cyber-enabled or cyber-dependent) is whether the interpreter of a transmission or message at any given point in the commission of a crime is a who (a person reading and interpreting a message) or a what (a machine resolving input) (Figure 4). The interpreter distinction in the communication process also influences the quality of the transmitted content. For example, it would be an unlikely choice for perpetrators to call an elderly person, proceed to read out a series of commands used in the Python programming language, and expect the call recipient's bank information in return. Similarly, submitting to a machine a lottery notification composed in broken English is unlikely to "persuade" the machine into further action. In either instance, however, the prerequisite for the possibility of a criminal offence under the crime-as-communication approach is an open channel between the sender and recipient, i.e. if a connection cannot be established, a crime(-as-communication) act cannot be completed.

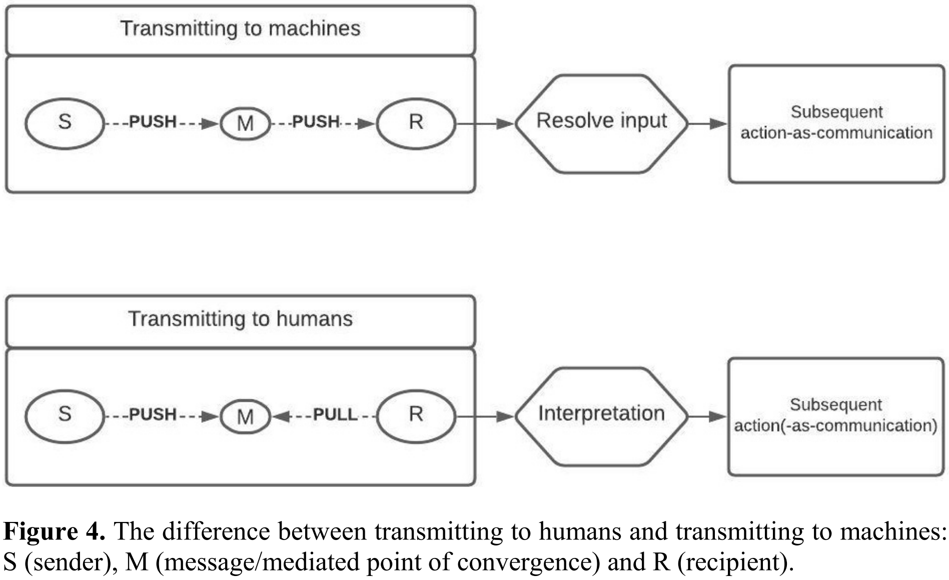

Figure 4. The difference between transmitting to humans and transmitting to machines: S (sender), M (message/mediated point of convergence) and R (recipient).

Hence, the success of the cybercriminal act from the perpetrators' perspective depends on the information interpreters that get (or do not get) in the way of crime acts based wholly or partly on communication. In crime acts targeting machines, this "interpretation" comes down to rule-based resolution of transmitted input. For instance, firewalls can be used to block certain incoming connections and running anti-virus software can detect known malicious code (CISA, 2019).

However, where a criminal offence requires action from a human recipient, such as in social engineering attacks, the success or failure of a crime attempt comes down to how the recipient interprets, understands and (potentially) acts on the basis of the received communication. Any individual characteristics of recipients, e.g. age, gender or education, are therefore relevant only insofar as these influence how recipients interpret incoming communications, both in terms of meaning (Hall, 1973) and legitimacy (Levine, 2019) (Figure 5).

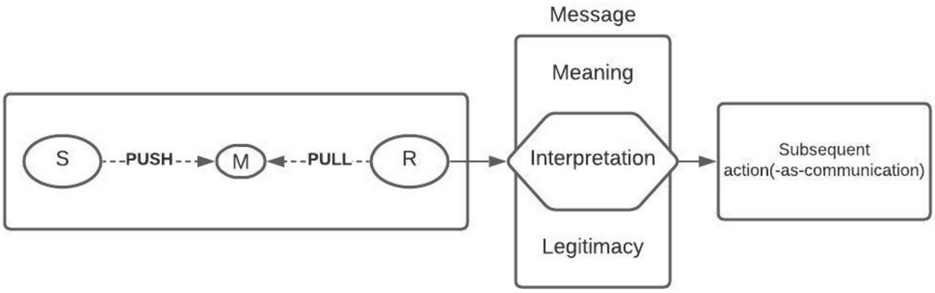

Figure 5. Recipients interpret received communications both in terms of meaning and legitimacy to decide on subsequent action.

An important quality of interpretations is their dynamic nature with respect to a single individual under different circumstances and with respect to different people experiencing the same circumstances. People may interpret messages differently depending on their personal circumstances (Lichtenberg, Stickley, & Paulson, 2013) and the social context in which the message is received (Rigotti & Rocci, 2006; **Study III**). This notion is in line with the dynamic understanding of vulnerability (see e.g. Kuran et al., 2020). For instance, while cultural and political structures, as well as historical forces, play a role in shaping societal vulnerability, it is important to recognise that groups (and, thus, individuals) are vulnerable in some and resilient in other contexts (Kuran et al., 2020) and that conditions of vulnerability are subject to change (Tierney, 2019 as cited in Kuran et al., 2020). For social engineering attacks, this means that vulnerability can be construed as the reason or reasons why some people with certain (combinations of) characteristics may, in some contexts or under certain circumstances, interpret as legitimate a communicated message that is fraudulent.

According to Levine (2019), if nothing in what is communicated triggers the recipient to suspect deception, interpreting the received message as truthful (legitimate) is people's default setting. Therefore, knowledge of diagnostically useful information relevant in instances of crime-as-communication requiring recipient interaction contributes to how recipients perform the interpretation process, i.e. how a recipient arrives at an interpretation of an incoming message's legitimacy (or lack thereof). As exemplified in **Study I**, **II** and **III**, a recipient may have to decide whether the World Health Organization actually sent them an email with an attachment containing health information, whether a sender really has obtained sensitive images depicting the recipient as well as whether the electricity will actually be shut off due to supposedly unpaid utilities bills etc. Elaborating on the information available to recipients from the messages and social context of social engineering attacks is therefore essential for preventing successful social engineering attacks. Such information must be general and robust enough to allow application in and across different communication media, message topics and social contexts (**Study I, II, III**). More importantly, this information ought to be easily understandable even without expert knowledge about cybercrime, i.e. to avoid overwhelming people with excessive details (Button & Cross, 2017).

Hence, raw input for interpretations comes from the observable choices that perpetrators make in producing and transmitting crime(-as-communication) messages and the relevant diagnostically useful information can, therefore, be obtained by explaining the function of such choices with respect to the interpretation process.

5.2 Choosing identities and choosing contexts

Elements relating to a sender's identity are the first items of information available to recipients, e.g. senders' email addresses, phone numbers and names. **Study I** showed that under normal circumstances, perpetrators choose to impersonate well-known companies or, alternatively, use made-up names or the mystique of anonymity. In combination with the body text of a message, all three of the aforementioned general choices of identity can be viewed as attempts at facilitating source credibility and authority (see e.g. Cialdini, 2009). The identity of the ailing widow, who offers up millions of dollars, is in line with previous research into perpetrator communications in which senders use up-front disclosure of personal information to establish a relationship with the recipient (see Carter, 2021). The same is evident from the threatening hacker message (**Study I**), because choosing the identity of an "anonymous hacker" means that the sender must disclose more information about their actions in the body of the message.

Conversely, impersonating a well-known brand or company lessens the explanatory workload required from the perpetrators. In other words, part of the persuasion is pre-established by the reputation of the company and/or brand, e.g. Amazon, PayPal or Facebook. This also means that message context is not created within the message itself (e.g. the widow's case), but the message is fitted to a pre-existing context. The latter was also clearly present in the findings of **Study II** and **III**, which showed a variety of ways that criminals exploited the salience of the COVID-19 pandemic when producing fraudulent messages. Furthermore, both the interpretive and constitutive dimensions of context (Rigotti & Rocci, 2006) were clearly present in scam messages relating to the COVID-19 pandemic (**Study II**, **III**). Common scam plots such as bogus offers for supermarket coupons took on a different meaning during the COVID-19 pandemic, as people lost jobs and suffered financially (**Study III**). In contrast, the pandemic context also allowed to circulate and gave credibility to messages that would not make much sense outside the COVID-19 context, e.g. notifications of fines imposed due to violations of lockdown rules (**Study II**, **III**).

Thus, the importance of message context cannot be underestimated. Combining the already established element of impersonation (Button & Cross, 2017) with message context provides new insights into the ways in which the content of fraudulent messages is framed (Figure 6). Where salient events or well-known company names are absent, scammers must rely more heavily on storytelling and disclosure of (bogus) information related to their made-up identity. Put differently, the scammers need to establish the chosen identity and storyline before they can effectively start making requests for action from the recipient, i.e. arrive at the core purpose of social engineering attacks. Hence, the exploitation of salient social contexts not only decreases criminals' workload, but is also likely to improve the meaningfulness of their communication as the social circumstances are "lived" by the sender and recipient alike (**Study III**).

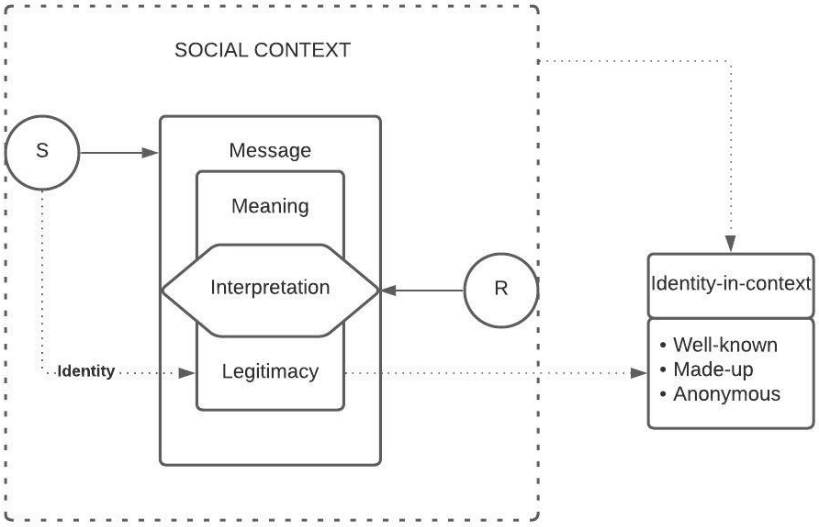

Figure 6. Sender's choice of identity-in-context for the purpose of framing the message and as an attempt to increase perceived legitimacy.

5.3 Persuasion or coercion?

The results of **Study III** showed that only 13.5% of scam messages (n= 140) used the coercive "Shock and Awe" approach to message production. To an extent, this finding is at odds with previous research, which has shown that, in comparison with the possibility of gain, people are more strongly influenced by the possibility of loss (Williams & Polage, 2019). However, newer research into phishing email detection has shown that people are more likely to dismiss threatening emails as not legitimate (McAlaney & Hills, 2020). In part, this can also explain why 86.5% (n= 900) scam messages used the gain-based "Good Samaritan" approach to message production. Since the ultimate goal of social engineering attacks is to get the recipient to perform actions that benefit the sender (Khonji, Iraqi, & Jones, 2013), early dismissal of the message due to lack of perceived legitimacy is opposite to what perpetrators are seeking from such communications (Figure 7).

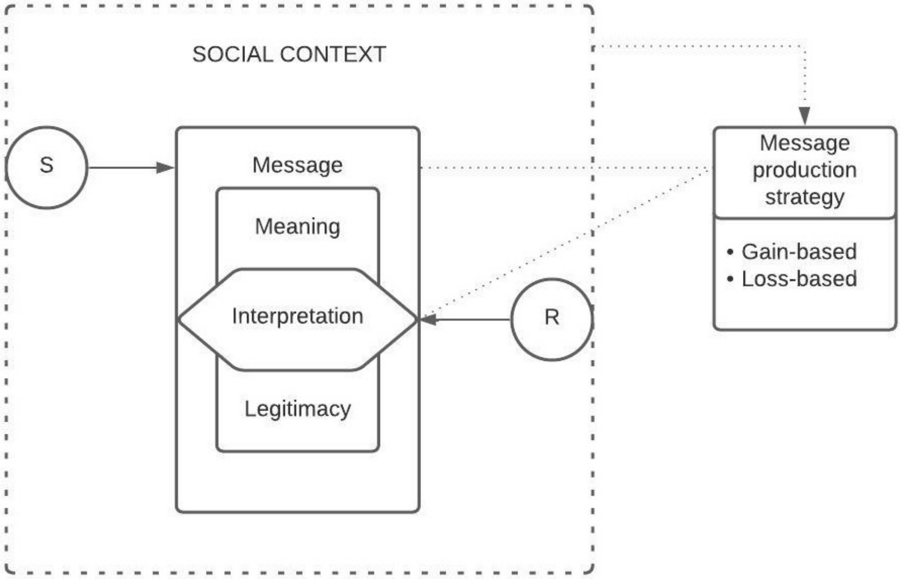

Figure 7. Sender's choice of message production strategy was influenced by the salient social context.

Furthermore, the social circumstances of the COVID-19 pandemic also provided scammers with more "legitimate" opportunities for offering relief rather than threatening loss. Scarcity of clear and credible information about the virus in the early weeks of the pandemic, subsequent shortages of personal protective equipment, and highly anticipated news of effective cures or even a vaccine all lent themselves as possible angles for a gain-based approach to producing fraudulent messages (**Study II**). While there were opportunities for a loss-based "Shock and Awe" approach, e.g. bogus fines and threats to infect the recipient's family members, these formed an obvious minority among the circulated scams (**Study III**).

Hence, scams that use salient social circumstances as context also reflect the specific needs that such social circumstances create or amplify – be it the need for information (health information), goods (groceries and essentials) or services (support services, healthcare). This includes whether criminals opt for a more persuasive or coercive approach in their general message production process. While previous research has shown that disaster situations, i.e. the social context of a hurricane, did entail adapted scam content during and after the event (Verma, Crane, & Gnawalli, 2018), the question remains as to what exactly determines the **salience threshold** for adapted scam message production.

Although Williams & Polage (2019) attempted to use the notion of a salient event in an experimental setting, the results were insignificant regarding the persuasiveness of the associated scam messages. As I also argued in **Study III**, this lack of significance could be because the social circumstances were not "lived" by the study participants. Thus, while an exact salience threshold is difficult to determine, the immediacy of important events or circumstances is significant for their persuasive use as social context in social engineering attacks (Verma, Crane, & Gnawalli, 2018; Steinmetz, Pimentel, & Roe, 2021; **Study II**, **III**). Still, what salient social circumstances add in terms of meaningfulness also take away from the longevity of such (thematically consistent) scams, i.e. when the social circumstances change or cease, the adapted scams lose their relevance and persuasive appeal.

Hence, I argue that the scam messages analysed in **Study I**, i.e. fraudulent messages not evidently based on specific social circumstances or an important event, are more resilient and thus likely to enjoy longer circulation. This means that specific social circumstances, which influence the content of scams, must be considered in addition to the more general (and robust) diagnostically useful information when detecting deception.

5.4 Generalised diagnostically useful information

The key piece of diagnostically useful information available to recipients of social engineering attack messages is the presence of a **request for action** (Step 2 in Figure 8**)**. Requests for action were present in all messages analysed in **Study I**, **II** and **III**. On one hand, the notion of requests for action as diagnostically useful information seems like a reiteration of the definition of social engineering attacks, i.e. the use of socially engineered messages the aim of which is to persuade the recipient to perform some action that benefits the sender (Khonji, Iraqi, & Jones, 2013). On the other hand, **Study I**, **II** and **III** showed how varied the types of requests found in social engineering attacks can be, e.g. ranging from widows offering millions of dollars to someone threatening to infect the recipient's family members with a virus. While the finding of a request for action was universal across the empirical studies **I, II, III** covered in my thesis, this finding needs to also be "translated" into an actionable and easy-to-understand piece of information that can aid recipients in detecting potential deception moving forward.

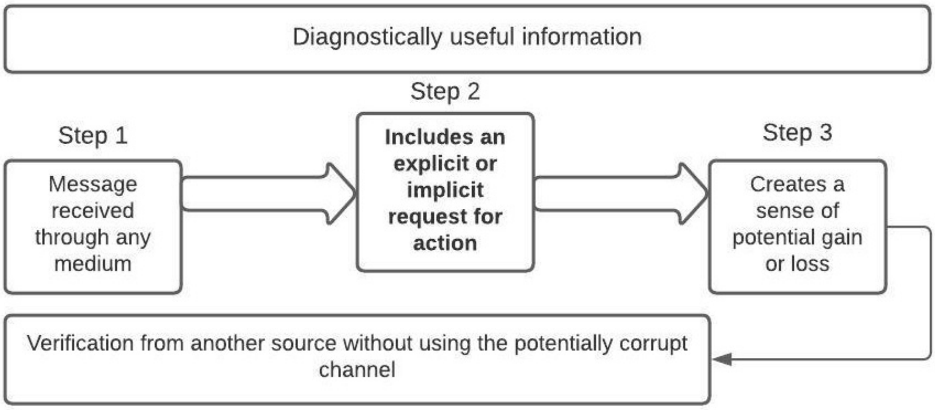
Diagnostically useful information
Step 1
Message received through any medium
Step 2
Includes an explicit or implicit request for action
Step 3
Creates a sense of potential gain or loss
Verification from another source without using the potentially corrupt channel

Figure 8. 3-step process that leads to verification of message legitimacy from another source without using the original (potentially corrupt) channel from which the message was received.

Thus, it is important to note that not all requests for action are **explicit**, e.g. directly ordering the recipient to make payments or clearly pointing out that the recipient should provide their bank details in a reply. **Implicit requests for action** also form an important category, in particular where the perpetrators only provide the recipient with a brief storyline or very little information, e.g. the case of bogus fines for violating lockdown rules (see **Study II, III**). Instances where the recipient is provided information, but not a clear directive, and a pathway to further action, e.g. a link in a text message or an attachment in an email, combine into implicit requests for action that must be understood similar to explicit requests.

The second piece of diagnostically useful information comes in the form understanding the variety of media or channels available to perpetrators for use in social engineering attacks. **Study III** showed that criminals use all major means for technology-mediated communication, i.e. emails, text messages, phone calls, social media posts and messages as well as websites, to perpetrate their attacks. Hence, recognising the importance of **open channels** between the perpetrator (sender) and target (recipient) must be, in the first instance, broader than merely focussing on one attack vector, e.g. phishing emails. This approach is particularly relevant due to the easy and (near-)immediate access that perpetrators currently have to recipients (**Study IV**). Furthermore, Button & Cross (2017) have emphasised that cybercrime prevention and awareness campaigns often confuse the target audience with excessive details about particular scams. Thus, as a practical directive, acknowledging the general importance of an open channel ought to come before focussing on specific channels and detailed scams perpetrated via such channels.

Additionally, the variety of media available to perpetrators for carrying out attacks as well as the need to verify suspicions of deception (Levine, 2019) lead to the idea of a **corrupt original channel**. For instance, in the case of the bogus COVID-19 fines, the initial socially engineered message came via text message (**Study III**). The verification process should therefore begin by choosing not to interact with the suspicious message in the original (presumably corrupt) channel, i.e. by responding to the phone number or email address from which the suspicious message was received. Secondly, the verification process ought to include checking the contents of the received message using another channel. It is important to note that the <u>medium</u> can remain the same, e.g. after receiving and hanging up a suspicious phone call, the recipient makes another call to legitimate source to verify the information received in the initial call.

The idea of the corrupt original channel therefore relates to avoiding (unnecessary) interaction with suspected scammers in the channel of the perpetrators' choosing (**Study III, IV**), which involves the medium and other aspects of message delivery such as the identity chosen by the criminal and thus associated with the original channel. However, in cases where the recipient suspects that a previously established channel, e.g. an email address through which a business partner communicates (**Study I**), has become corrupted, then the verification process must rely on a different medium. Hence, the verification process should involve a different source of information where the medium used for verification remains the same, and a different medium where the verification relates to whether a source would actually send the message received in another channel.

Furthermore, as I already mentioned above (see 4.3) and argued in **Study IV**, the protective role of third persons is only secondary in the verification process. In other words, where the initial recipient 1) does not suspect deception in the received message and 2) does not attempt to verify the legitimacy of the message, third persons are unlikely to even become involved in the verification process. That is because third persons are unlikely to have immediate and direct access to the same channel through which the recipient was contacted. Thus, the detection of deception resides, in the first instance, with the initial recipient, while external sources, including third persons, become relevant only in the course of the subsequent verification process.

Therefore, it is imperative for recipients to know that the typical scam message primarily relies on either a gain-based or a loss-based approach to producing message content (**Study II**, **III**). Hence, the third piece of diagnostically useful information relates specifically to the content of social engineering attacks, i.e. whether the request for action connects to 1) promised gains or 2) threatened losses. Alsharnouby, Alaca, & Chiasson (2015) found that people pay more attention to the content of the message rather than its technical details.

Furthermore, recent research into phishing detection that used eye-tracking technology, i.e. to ascertain which parts of the email recipients read first, suggests that recipients look for threatening content and urgency cues before aspects such as misspellings (McAlaney & Hill, 2020). Thus, focussing on content – i.e. "what is said" (Levine, 2019) – is crucial if we consider that other media, e.g. phone calls, used by perpetrators to carry out social engineering attacks do not necessarily offer the same level of technical detail for scrutiny by the recipients.

With respect to the amount of content available for scrutiny, it is important to note that the scams analysed in **Study I, II** and **III** were brief in nature, i.e. the (explicit or implicit) request for action was revealed in the first message, which is not always the case. For instance, in longer-running romance scams (Carter, 2021), the perpetrator first establishes contact with the recipient, builds trust over time and only then, at some point, levies the illicit request for action.

In a way, such longer-running scams are an advanced version of the scams analysed in **Study I** where contact was made by an initially unknown person and the request for action was presented promptly. The more advanced nature of scams that hold off on the request for action comes from the fact that the delay allows to build context where there first was none. For instance, in the widow example in **Study I**, the perpetrators attempted to disclose as much bogus information about the made-up sender identity as possible in one message, while romance scams as described by Carter (2021), build this disclosure over time, thus allowing for the creation of personal context. However, as **Study II** and **III** showed, the presence of salient social contexts makes it possible for scammers, who are willing to quickly adapt message content, to skip the context-building portion of fraudulent communications and simply draw on the circumstances already affecting recipients.

Additionally, being aware that perpetrators may use specific influencing techniques (Cialdini, 2009; **Study III**) in their messages can further help recipients detect deception. Knowing the basics of the six principles of persuasion proposed by Cialdini (2009), i.e. the principles of authority, scarcity, liking, social proof, reciprocity and consistency, is an important tool for a more nuanced detection of influencing, including where this is carried out for fraudulent purposes. However, recognising the "Good Samaritan" and "Shock and Awe" approaches as these appear in received messages (**Study II**) functions as a shorthand for the more detailed principles of persuasion (**Study III**). This is because the combination of presenting a request for action and leading the recipient to that action by promising gains or threatening loss forms the core of every technology-mediated scam. Nevertheless, the way in which Cialdini's principles are used in other everyday contexts such as shopping and commerce (see e.g. Adaji, Oyibo, & Vassileva, 2020) could prove to be a practical course for also introducing their application in social engineering attacks.

5.5 Limitations and future research

The research presented in my thesis has some limitations. Following Brantingham & Brantingham (1981), there are four dimensions to understanding crime: the legal dimension, the offender dimension, the victim dimension and the place (situational) dimension. In the three empirical works and the theoretical criticism that form the basis of my thesis, the legal dimension was addressed in **Study I**, while all studies dealt with the situational dimension, i.e. the mediated point of convergence between the perpetrator (sender) and target (recipient). This means that primary attention was not given to the offender and victim dimensions, which is both the main limitation of my thesis as well as the pathway for future research with respect to the crime-as-communication approach. The message element of crime-as-communication and service providers that function as the mediated point of convergence element are equally important moving forward. The core elements of the crime-as-communication approach are the basic building blocks in a much wider ecosystem of technology-mediated social engineering attacks (see Figure 9).

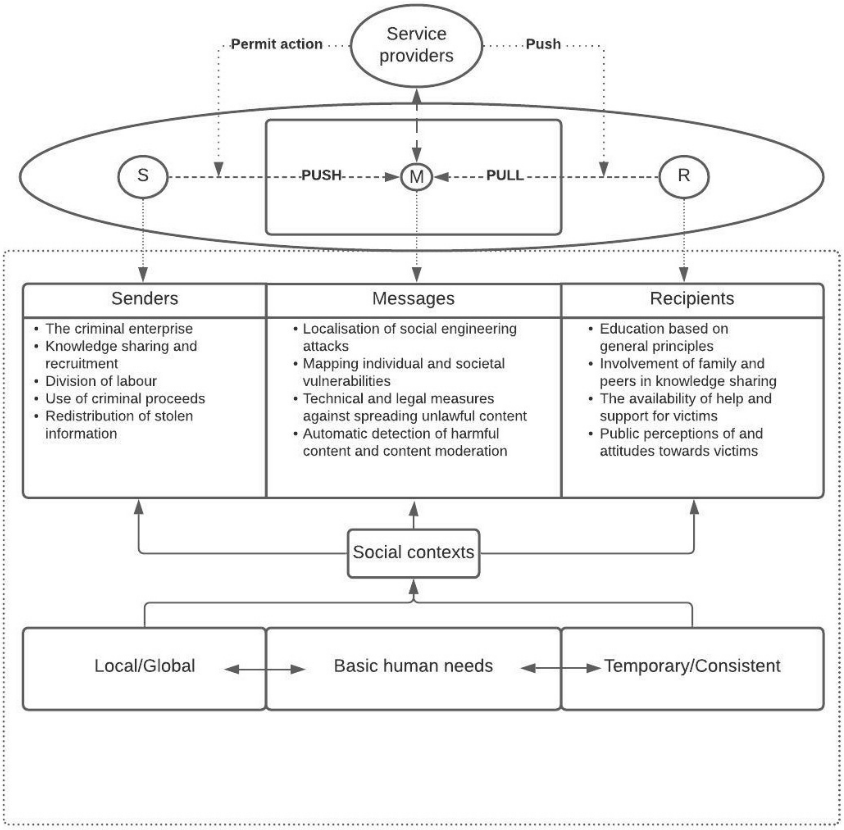
Service providers
Permit action
Push
S
PUSH
M
PULL
R
Senders
• The criminal enterprise
• Knowledge sharing and recruitment
• Division of labour
• Use of criminal proceeds
• Redistribution of stolen information
Messages
• Localisation of social engineering attacks
• Mapping individual and societal vulnerabilities
• Technical and legal measures against spreading unlawful content
• Automatic detection of harmful content and content moderation
Recipients
• Education based on general principles
• Involvement of family and peers in knowledge sharing
• The availability of help and support for victims
• Public perceptions of and attitudes towards victims
Social contexts
Local/Global
Basic human needs
Temporary/Consistent

Figure 9. The social engineering attack ecosystem: core elements, social contexts and research foci.

5.5.1 Senders: techniques and division of labour

Criminals are known to converge in online venues to seek out co-offenders, request and exchange information and make plans about carrying out illegal activities (Leukfeldt, Kleemans, & Stol, 2017). Recent reports note that the cybercriminal division of labour is very complex and specialised (Proofpoint, 2021), i.e. who provides access to the target (opens the "channel"), who produces the necessary malware (composes the "message") and who carries out the attack. The first of the aforementioned – the so-called initial access brokers - are the likeliest to use social engineering attacks to gain access to companies' networks. Previous works (see Zielinska, Welk, Mayhorn, & Murphy-Hill, 2016) and my thesis (**Study III**) have shown that the use of specific influencing techniques can be detected from scam messages. Nevertheless, future research could further address whether the use of such techniques is a conscious choice by offenders based on their knowledge of these techniques or does the use rely more on information obtained from online venues, e.g. hacker forums, and the techniques are implemented merely in a formulaic manner. What is more, unlike Cialdini's (2009) principles of persuasion, which are not all equally present in every scam message, future research could certainly address perpetrators' choices regarding the use of "Good Samaritan" and "Shock and Awe" (**Study II**) approaches to message production. In particular, research can address in more detail whether these choices in message production are influenced by the social circumstances, assumed psychological vulnerabilities of the targets or the nature of topics and themes that the perpetrators seek to exploit.

5.5.2 Messages and service providers

Study III showed that some criminals adapt the scams being circulated to reflect salient social circumstances relatively quickly. Nevertheless, this process is not always without delays, which have been explained by insufficient English skills of the scammers (see EC News Desk, 2020 in **Study III**). However, recent developments (see Gendre, 2021) further suggest that scammers are widening the so-called language of cybercrime by more frequently including languages other than English into the production of scam messages. This causes severe problems for automated detection systems, which are language dependent (Jain & Gupta, 2021), and further elevates the importance of deception detection by recipients. Some research into phishing in native languages exists (Kävrestad, Pettersson, & Nohlberg, 2020), but social engineering attacks perpetrated in recipients' native languages require more research attention moving forward. In particular, the focus could be on whether social engineering attacks carried out in recipients' native languages are interpreted differently by the recipients, e.g. in terms of perceived source credibility or deception detection, and what role is played by the recipients' own language skills.

With respect to service providers (see Figure 9), my thesis focused on social engineering attacks that are often carried out using media the operators of which are subject to strict message secrecy regulations, i.e. telecommunications companies. However, the push-and-pull of content/messages between senders and recipients can just as well be played out using other points of mediated convergence, e.g. social media platforms or online forums. A crucial question with respect to any point of mediated convergence is whether and, if yes, to what extent are the relevant operators involved in pushing content to recipients. In particular, this concerns social media platforms and forums that delay removal of disinformation or harmful content (see e.g. Bond, 2020; Kikerpill, Siibak, & Valli, 2021), i.e. why senders are allowed to "push" such content in the first place, why the mediated point of convergence delays removal or takedown action, and why checks on who is able to "pull" the content are relaxed.

5.5.3 Recipients: education and aftercare

Finally, my empirical works analysed and provided guidelines on the principle mechanics of social engineering attacks – i.e. diagnostically useful information – based on the content and context of the latter, which does not provide primary accounts of sender intentions and recipient reactions. However, the diagnostically useful information analysed in my thesis can be used in future educational programmes meant for people without expert interest in or knowledge about cybercrime. Key elements (Figure 8), i.e. knowing that scams are not tied to specific media; that scammers always, at some point, present an implicit or explicit request for action; and that this request is always broadly based on either promising gains or threatening losses, are easier to convey than the myriad of plotlines used in scams, which can confuse the target audience (Button & Cross, 2017) and should thus only serve as examples of the aforementioned principles. Therefore, future research could focus more on how to convey these principles to different audiences, e.g. age groups, rather than continuing the practice of only or primarily conveying excessive details about specific scams. Research could also further address questions related to those who could potentially convey such knowledge, e.g. parents and teachers to children, relatives or social workers to the elderly, relevant public authorities to residents etc.

The so-called "learning the hard way" side of social engineering attacks, i.e. the psychological and financial impacts from becoming a victim of such attacks, must also be front and centre in public discussions and publicly funded research. As previous works (Button & Cross, 2017; Carter, 2021) and my thesis have shown (see also Figure 9), perpetrators of social engineering attacks most often simply exploit the humanity of their targets. In other words, even without the presence of temporary or periodic salient events, criminals are incredibly adept at exploiting recipients' basic human needs to distort the perception of the latter regarding incoming communications. Further mapping of the ways in which recipients' basic human needs are exploited in various types of social engineering attacks is needed, in particular with respect to how these exploits relate to the dynamic understanding of vulnerability on the individual and societal levels, and in local and global contexts. This includes questions regarding the help and support that is available to victims of cybercrime as well as issues relating to the public perception and attitude towards cybercrime victims, e.g. victims of fraud, bullying, harassment, stalking and other forms of abuse.

CONCLUSION

My thesis has focused on the diagnostically useful information that is available to recipients from the content and context of social engineering attacks. Following from the premise that every activity in technology-mediated environments is rooted in and dependent upon communication, I have previously provided the theoretical background, results and discussion for establishing the **crime-as-communication** approach. Understanding the strategies of persuasion and coercion that perpetrators rely on to influence recipients into taking action that is harmful to them and knowing how these strategies appear in socially engineered messages is the key to detecting and countering incoming crime(-as-communication) attempts. Thus, in this Chapter, I will conclude my thesis by answering the research questions presented earlier (see 2.4.1):

What diagnostically useful information is available to a recipient from the content and context of social engineering attacks?

a. How can conventional crimes defined under criminal law inform the communicative aspects of criminal activity in mediated form?

The offences of fraud and extortion provide important input for understanding crimes-as-communication as these offences rely on influencing the crime targets' (message recipients') will to act (**Study I**) and require action from the target, which are the hallmarks of social engineering attacks. The general approach taken by senders, i.e. persuasion or coercion, ought to be considered more important than the specific plotlines used in social engineering attacks.

b. Which general communicative approaches appear in the production of messages used in social engineering attacks?

I developed and analysed fraud- and extortion-type communications as these appeared in phishing emails (**Study I**) and used them as input for developing the gain-based "Good Samaritan" and the loss-based "Shock and Awe" approaches to message production (**Study II**). **Study III** quantified the appearance of the "Good Samaritan" and "Shock and Awe" approaches to message production over the first four months of the COVID-19 pandemic (January-April 2020), showing that 900 of 1040 scams analysed (86.5%) used the gain-based and 140 (13.5%) used the loss-based approach. The presence of (one of) the two approaches is important as diagnostically useful information.

c. What role does impersonation play in social engineering attacks?

Impersonation, including "spoofing", plays a key role in all social engineering attacks even where the perpetrators choose to create non-existent identities or contact recipients anonymously (**Study I**). **Study II** and **III** showed that, salient social circumstances permitting, perpetrators adapt impersonation choices to carry out contextually fitting social engineering attacks.

d. Which media are used by criminals to perpetrate social engineering attacks?

Study III showed that while email was the most popular medium for scams during the first four months of the COVID-19 pandemic (53.5%), phone calls (13.6%), text messages (12.6%), bogus websites (11.3%) and social media messages and posts (3.1%) also played a significant role. This variety of media allows to focus on the presence of open channels between senders and recipients rather than specific media, which is key in understanding the ease with which criminals can establish contact with their targets (**Study IV**). With respect to providing diagnostically useful information, social engineering attacks should be considered as medium-independent.

e. What topics and themes do criminals rely on in social engineering attacks?

Study I showed that, under normal circumstances, criminals use generic topics to capture as wide an audience as possible and that the use of similar or identical themes can occur in waves. **Study II** and **III** indicated that given specific and salient social circumstances, criminals are adept at adapting the topics, themes and content of their scam messages to fit the relevant social context. Hence, given the relevant opportunity, some scammers adapt the topics and content of their scam messages in a way that reflects sudden changes in social circumstances, which suggests that perpetrators consider such adaptations to be more profitable than the generic topics used under normal circumstances.

f. Which specific social-psychological influencing techniques are employed in social engineering attacks?

Study III used Cialdini’s six principles of persuasion as an additional layer of qualitative content analysis and found empirical support for all six: authority, scarcity, liking, social proof, consistency and reciprocity. The fact that all six principles found empirical support further emphasises the extent to which cybercriminals go to influence crime-as-communication recipients. Knowledge about the use of Cialdini’s principles in other contexts, e.g. commerce, could provide a practical route for teaching about their use in social engineering attacks.

g. What role does social context play in social engineering attacks?

Study I suggested that perpetrators put more effort into crafting bogus stories where salient social circumstances were absent, i.e. the context for the interaction was established within the message itself. **Study II** and **III** showed that sudden changes in salient social circumstances such as the COVID-19 pandemic, motivate perpetrators to adjust the content of scams in order to reflect said circumstances. This was reflected in the themes, impersonated identities and choice of approach in scam message production.

How does the crime-as-communication approach contribute to current criminological thought?

Rather than attempting to fit a ‘terrestrial’ (or conventional) understanding of action and convergence onto acts of cybercrime, the crime-as-communication approach instead builds a communication-based foundation for understanding action in technology-mediated environments. By making explicit the communication processes that actually constitute cybercriminal activity, the crime-as-communication approach provides a detailed and principled way of analysing technology-mediated crime specifically, and action in technology-mediated environments more broadly. Moreover, the crime-as-communication approach emphasises the role that open channels and the interpretation of messages play in avoiding becoming a victim of cybercrime. Where an open channel exists, recipient characteristics (humans) or properties (machines) are relevant under the crime-as-communication approach insofar as these influence interpretations of incoming communications, i.e. detecting harmful or criminal input from all received input.

REFERENCES

Adaji, I., Oyibo, K., & Vassileva, J. (2020). E-commerce shopping motivation and the influence of persuasive strategies. *Frontiers in Artificial Intelligence*, 3(67). https://doi.org/10.3389/frai.2020.00067

Addams, J. (1914). The larger aspects of the woman's movement. *The ANNALS of the American Academy of Political and Social Science*, 56(1), 1–8. https://doi.org/10.1177/000271621405600101

Alexander, J., & Schmidt, J. K. H. W. (1996). Social engineering: genealogy of a concept. In A. Podgorecki, J. Alexander, & R. Shields (Eds), *Social Engineering* (pp. 1–19). McGill-Queen's University Press.

Alsharnouby, M., Alaca, F., & Chiasson, S. (2015). Why phishing still works: user strategies for combating phishing attacks. *International Journal of Human-Computer Studies*, 82, 69–82.

Andresen, M. A. (2010). The place of environmental criminology within criminological thought. In M. A. Andresen, P. J. Brantingham, & J. B. Kinney (Eds), *Classics in Environmental Criminology* (pp. 5–28). Burnaby: Simon Fraser University Publications.

Armstrong, M. E., Jones, K. S., & Namin, A. S. (2021). How perceptions of caller honesty vary during vishing attacks that include highly sensitive or seemingly innocuous requests. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. https://doi.org/10.1177/00187208211012818

Aziz, A. (2020). Facebook ad boycott campaign 'Stop Hate for Profit' gathers momentum and scale: inside the movement for change. Accessed 28.06.2021, https://www.forbes.com/sites/afdhelaziz/2020/06/24/facebook-ad-boycott-campaign-stop-hate-for-profit-gathers-momentum-and-scale-inside-the-movement-for-change/?sh=7f67a7af1668

Bax, S., McGill, T., & Hobbs, V. (2021). Maladaptive behaviour in response to email phishing threats: the roles of rewards and response costs. *Computers & Security*, 106. https://doi.org/10.1016/j.cose.2021.102278

Bond, S. (2020). Civil rights groups say if Facebook won't act on election misinformation, they will. Accessed 29.06.2021, https://www.npr.org/2020/09/25/916782712/civil-rights-groups-say-if-facebook-wont-act-on-election-misinformation-they-wil?t=1624731537683

Bossler, A. M., & Holt, T. J. (2009). On-line activities, guardianship, and malware infection: an examination of routine activity theory. *International Journal of Cyber Criminology*, 3(1), 400–420.

Brantingham, P. J., & Brantingham, P. L. (1981). Introduction: The dimensions of crime. In P. J. Brantingham, & P. L. Brantingham (Eds), *Environmental Criminology* (pp. 7–26). Prospect Heights, IL: Waveland Press.

Brody, R. G., Kern, S., & Ogunade, K. (2020). An insider's look at the rise of Nigerian 419 scams. *Journal of Financial Crime*. https://doi.org/10.1108/JFC-12-2019-0162

Brown, B. (2021). Responsibilization and recovery: shifting responsibilities on the journey through mental health care to social engagement. *Social Theory & Health*, 19, 92–109. https://doi.org/10.1057/s41285-019-00097-x

Buil-Gil, D., Lord, N., & Barrett, E. (2021). The dynamics of business, cybersecurity and cyber-victimization: foregrounding the internal guardian in prevention. *Victims & Offenders*, 16(3), 286–315. https://doi.org/10.1080/15564886.2020.1814468

Burchell, G. (1996). Liberal government and techniques of the self. In A. Barry, T Osborne, & N. Rose (Eds), *Foucault and Political Reason: Liberalism, Neo-liberalism, and Rationalities of Government* (pp. 19–36). Chicago: University of Chicago Press.

Button, M., Lewis, C., & Tapley, J. (2009). Fraud typologies and the victims of fraud: literature review. London: National Fraud Authority.

Button, M., & Cross, C. (2017). *Cyber Frauds, Scams and Their Victims*. Oxon: Routledge.

Capeller, W. (2001). Not such a neat net: some comments on virtual criminality. *Social & Legal Studies*, 10(2), 229–242. https://doi.org/10.1177/a017404

Carter, E. (2015). The anatomy of written scam communications: an empirical analysis. *Crime, Media, Culture: An International Journal*, 11(2), 89–103. https://doi.org/10.1177/1741659015572310

Carter, E. (2021). Distort, extort, deceive and exploit: exploring the inner workings of a romance fraud. *British Journal of Criminology*, 61(2), 283–302. https://doi.org/10.1093/bjc/azaa072

Chiluwa, I. (2019). “Congratulations, Your Email Account Has Won You €1,000,000”: analyzing the discourse structures of scam emails. In T. Docan-Morgan (Ed), *The Palgrave Handbook of Deceptive Communication* (pp. 897–912). Springer.

Cialdini, R. B. (2009). *Influence: Science and Practice*. Boston: Pearson Education.

CISA. (2019). Security tip (ST18-004): protecting against malicious code. Accessed 29.06.2021, https://us-cert.cisa.gov/ncas/tips/ST18-271

Clarke, R. V. (2013). Crime science. In E. McLaughlin, & T. Newburn (Eds), *The SAGE Handbook of Criminological Theory* (pp. 271–283). Thousand Oaks, CA: SAGE Publications, Inc.

Collie, C. J. R., & Greene, K, S. (2016). Stranger child abduction and guardianship: accompaniment and surveillance in attempted and completed cases. *Crime Prevention and Community Safety*, 18, 284–308. https://doi.org/10.1057/s41300-016-0002-3

Conroy, M. (2017). *Nazi Eugenics: Precursors, Policy, Aftermath*. Stuttgart: *ibidem* Press.

Denery, D. G. (2015). *The Devil Wins: A History of Lying from the Garden of Eden to the Enlightenment*. Princeton, NJ: Princeton University Press.

De Kimpe, L., Walrave, M., Verdegem, P., & Ponnet, K. (2021). What we think we know about cybersecurity: an investigation of the relationships between perceived knowledge, internet trust, and protection motivation in a cybercrime context. *Behavior & Information Technology*. https://doi.org/10.1080/0144929X.2021.1905066

Diamond, B., & Bachmann, M. (2015). Out of the beta phase: obstacles, challenges, and promising paths in the study of cyber criminology. *International Journal of Cyber Criminology*, 9, 24–34. http://doi.org/10.5281/zenodo.22196

Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union European Commission (19.07.2016). *Official Journal of the European Union*, L 194, pp. 1–30. Accessed 29.05.2021, http://data.europa.eu/eli/dir/2016/1148/oj

Douglass, R. (2020). Bernard Mandeville on the use and abuse of hypocrisy. *Political Studies*. https://doi.org/10.1177/0032321720972617

Drake, C. (2021). How businesses can protect their brands and bottom lines against damage from illegal robocalls. Accessed 20.06.2021, https://www.helpnetsecurity.com/2021/05/28/illegal-robocalls/

EC News Desk. (2020). SophosLabs tracks significant uptick in COVID-19 scams and phishing attacks. Accessed 29.05.2021, https://www.ec-mea.com/sophoslabs-tracks-significant-uptick-in-covid-19-scams-and-phishing-attacks/

Eck, J. E. (2003) Police problems: the complexity of problem theory, research and evaluation. In J. Knutsson (Ed), *Problem-Oriented Policing: From Innovation to Mainstream*. *Crime Prevention Studies, Vol. 15* (pp. 67–102). Monsey, NY: Criminal Justice Press.

Eck, J. E., & Madensen, T. D. (2015). Meaningfully and artfully reinterpreting crime for useful science: an essay on the value of building with simple theory. In M. A. Andresen, & G. Farrell (Eds), *The Criminal Act: The Role and Influence of Routine Activity Theory* (pp. 5–18). Hampshire: Palgrave Macmillan.

European Commission. (2013). Impact Assessment accompanying the document Proposal for a Directive of the European Parliament and of the Council Concerning measures to ensure a high level of network and information security across the Union. Commission Staff Working Document, SWD/2013/032 final. Accessed 29.05.2021, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=SWD:2013:032:FIN

European Commission. (2017). *Resilience, Deterrence and Defence: Building strong cybersecurity for the EU*. Joint Communication to the European Parliament and the Council, JOIN/2017/0450 final. Accessed 29.05.2021, https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:52017JC0450

European Commission. (2020). *Europeans' Attitudes Towards Cyber Security*. Special Eurobarometer 499, Report.

Eurostat. (2019). Hours of work – annual statistics. Accessed 30.05.2021, https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Hours_of_work_-_annual_statistics#Working_hours_for_employees

Ezzy, D. (2002). *Qualitative Analysis: Practice and Innovation*. London: Routledge.

Fein, M. L. (2001). Social engineering in context: some observations on Turner. *Sociological Practice*, 3(2), 121–125.

Felson, M. (2008). Routine activity approach. In R. Wortley, & L. Mazerolle (Eds), *Environmental Criminology and Crime Analysis* (pp. 70–77). Portland, Oregon: Willan Publishing.

Felson, M. (2014). Interview recording (interviewed by B. Dooley). Oral History of Criminology. Accessed 30.05.2021, https://youtu.be/96QeOVGlttI

Felson, M. (2019). Choosing crime in everyday life: routine activity and rational choice theories. In J. R. Lilly, F. T. Cullen, & R. A. Ball (Eds), *Criminological Theory: Context and Consequences* (pp. 612–656). Thousand Oaks, CA: SAGE Publications, Inc.

Felson, M., & Boba, R. (2010). *Crime and Everyday Life*. Thousand Oaks, CA: SAGE Publications, Inc.

Felson, M., & Clarke, R. V. (1998). *Opportunity Makes the Thief: Practical Theory for Crime Prevention.* Police Research Series, Paper 98. London: Home Office.

Felson, M., & Clarke, R. V. (2011). The origins of the routine activity approach and situational crime prevention. In F. T. Cullen, C. L. Johnson, A. J. Myer, & F. Adler (Eds) *The Origins of American Criminolgoy (Advances in Criminological Theory, Volume 16)* (pp. 245–260). New York: Routledge.

Felson, M., & Cohen, L. E. (1979). Social change and crime rate trends: a routine activity approach. *American Sociological Review*, 4(44), 588–608.

Flinders, K. (2020). Banks report surge in impersonation scams. Accessed 29.06.2021, https://www.computerweekly.com/news/252489106/Banks-report-surge-in-impersonation-scams

Garland, D. (1996). The limits of the sovereign state: strategies of crime control in contemporary society. *British Journal of Criminology*, 36(4), 445–471.

Gendre, A. (2021, February 26). The rise of non-English language spear phishing emails. *Help Net Security*. Accessed 30.05.2021, https://www.helpnetsecurity.com/2021/02/26/non-english-spear-phishing-emails/

General Part of the Civil Code Act (1.04.2021). *Riigi Teataja I*. Accessed 29.05.2021, https://www.riigiteataja.ee/en/eli/501042021006/consolide

Grabosky, P. N. (2001). Virtual criminality: old wine in new bottles? *Social & Legal Studies*, 10(2), 243–249. https://doi.org/10.1177/a017405

Grabosky, P. N., & Smith, R. (2001). Telecommunication fraud in the digital age: the convergence of technologies. In D. Wall (Ed), *Crime and the Internet* (pp. 29–43). London: Routledge.

Gray, J. (1842). *An Efficient Remedy for the Distress of Nations*. London: Longman, Brown, Green and Longmans.

Greene, K., Steves, M., Theofanos, M., & Kostick, J. (2018). User context: an explanatory variable in phishing susceptibility. *Proceedings of the Network and Distributed Systems Security (NDSS) Symposium, San Diego, CA* (pp. 1–14). https://doi.org/10.14722/usec.2018.23016

Gunaratne, C., Baral, N., Rand, W., Garibay, I., Jayalath, C., & Senevirathna, C. (2020). The effects of information overload on online conversation dynamics. *Computational and Mathematical Organization Theory*, 26, 255–276. https://doi.org/10.1007/s10588-020-09314-9

Hadnagy, C. (2018). *Social Engineering: The Science of Human Hacking*. Indianapolis, IN: Wiley.

Hall, S. (1973). Encoding and decoding in television discourse. Discussion Paper. Birmingham: University of Birmingham. Accessed 30.05.2021, http://epapers.bham.ac.uk/2962/1/Hall%2C_1973%2C_Encoding_and_Decoding_in_the_Television_Discourse.pdf

Hansson, S., Orru, K., Torpan, S., Bäck, A., Kazemekaityte, A., Meyer, S. F., Ludvigsen, J., Savadori, L., Galvagni, A., & Pigrée, A. (2021). COVID-19 information disorder: six types of harmful information during the pandemic in Europe. *Journal of Risk Research*, 24(3), 380–393. https://doi.org/10.1080/13669877.2020.1871058

Hatfield, J. M. (2018). Social engineering in cybersecurity: the evolution of a concept. *Computers & Security*, 73, 102–113. https://doi.org/10.1016/j.cose.2017.10.008

Hawley, A. (1950). *Human Ecology: A Theory of Community Structure*. New York: Ronald.

Help Net Security. (2021, March 24). Remote workers admit to playing a significant part in increasing their company's cybersecurity risks. Accessed 29.05.2021, https://www.helpnetsecurity.com/2021/03/24/remote-workers-cybersecurity-risks/

Henry, N., & Powell, A. (2014). The dark side of the virtual world: towards a digital sexual ethics. In N. Henry, & A. Powell (Eds), *Preventing Sexual Violence: Interdisciplinary Approaches to Overcoming a Rape Culture* (pp. 84–104). Hampshire: Palgrave Macmillan.

Heyman, G. D., Luu, D. H., & Lee, K. (2009). Parenting by lying. *Journal of Moral Education*, 38(3), 353–369. doi: 10.1080/03057240903101630

Hollis, M. E., Felson, M., & Welsh, B. C. (2013). The capable guardian in routine activities theory: a theoretical and conceptual reappraisal. *Crime Prevention and Community Safety*, 15(1), 65–79.

Hollis-Peel, M. E., Reynald, D. M., Van Bavel, M., Elffers, H., & Welsh, B. C. (2011). Guardianship for crime prevention: a critical review of literature. *Crime, Law and Social Change*, 56(1), 53–70.

Holt, T. J., & Bossler, A. M. (2008). Examining the applicability of lifestyle-routine activities theory for cybercrime victimization. *Deviant Behavior*, 30(1), 1–25. https://doi.org/10.1080/01639620701876577

Hong, J. (2012). The state of phishing attacks. *Communications of the ACM*, 55(1). doi:10.1145/2063176.2063197

Hutchings, A., & Hayes, H. (2009). Routine activity theory and phishing victimisation: who gets caught in the net? *Current Issues in Criminal Justice*, 20(3), 433–452.

IBM Global Technology Services. (2014). IBM security services 2014 cyber security intelligence index.

Jacques, S. (2014). The quantitative-qualitative divide in criminology: a theory of ideas' importance, attractiveness, and publication. *Theoretical Criminology*, 18(3), 317–334. https://doi.org/10.1177/1362480613519467

Jagatic, T. N., Johnson, N. A., Jakobsson, M., & Mencer, F. (2007). Social phishing. *Communications of the ACM*, 50(10), 94–100. https://doi.org/10.1145/1290958.1290968

Jain, A. K., & Gupta, B. B. (2021). A survey of phishing attack techniques, defence mechanisms and open research challenges. *Enterprise Information Systems*. doi: 10.1080/17517575.2021.1896786

Jakobsson, M. (2007). The Human Factor in Phishing. *Privacy & Security of Consumer Information*.

Jõgi, M. (2012). *Verbaalselt täidetavad süüteokoosseisud Eesti karistusõiguses* [Verbally committable offences in Estonian criminal law]. Master's thesis. University of Tartu, School of Law.

Kävrestad, J., Pettersson, R., & Nohlberg, M. (2020). The language effect in phishing susceptibility. In P. Bednar, A. Nolte, M. Rajanen, H. V. Hult, A. S. Islind & F. Pigni (Eds), *Proceedings of the 6th International Workshop on Socio-Technical Perspective in IS Development (STIPS 2020)* (pp. 162–167). CEUR-WS.

Khonji, M., Iraqi, Y., & Jones, A. (2013). Phishing detection: A literature survey. *IEEE Communications Surveys & Tutorials*, 15(4), 2091–2121.

Kikerpill, K., Siibak, A., & Valli, S. (2021). Dealing with deepfakes: Reddit, online content moderation, and situational crime prevention. In J. B. Wiest (Ed), *Theorizing Criminality and Policing in the Digital Media Age* (pp. 25–45). Emerald Insight. doi: 10.1108/S2050–206020210000020008

Kim, D., & Kim, J. H. (2013). Understanding persuasive elements in phishing e-mails: a categorical content and semantic network analysis. *Online Information Review*, 37(6), 835–850. https://doi.org/10.1108/OIR-03-2012-0037

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology*. Thousand Oaks, CA: SAGE Publications, Inc.

Kritzinger, E., & Von Solms, S. H. (2010). Cyber security for home users: a new way of protection through awareness enforcement. *Computers & Security*, 29(8), 840–847.

Kuran, C. H. A., Morsut, C., Kruke, B. I., Krüger, M., Segnestam, L., Orru, K., Nævestad, T. O., Airola, M., Keränen, J., Gabel, F., Hansson, S., & Torpan, S. (2020). Vulnerability and vulnerable groups from an intersectionality perspective. *International Journal of Disaster Risk Reduction*, 50, 101826. https://doi.org/10.1016/j.ijdrr.2020.101826

Leukfeldt, E. R., Kleemans, E. R., & Stol, W. P. (2017). Cybercriminal networks, social ties and online forums: social ties versus digital ties within phishing and malware networks. *British Journal of Criminology*, 57(3), 704–722. https://doi.org/10.1093/bjc/azw009

Leukfeldt, E. R., & Yar, M. (2016). Applying routine activity theory to cybercrime: a theoretical and empirical analysis. *Deviant Behavior*, 37(3), 263–280. https://doi.org/10.1080/01639625.2015.1012409

Levine, T. R. (2019). *Duped: Truth-Default Theory and the Social Science of Lying and Deception*. Tuscaloosa, AL: University of Alabama Press.

Lichtenberg, P. E., Stickney, L., & Paulson, D. (2013). Is psychological vulnerability related to the experience of fraud in older adults? *Clinical Gerontologist: The Journal of Aging and Mental Health*, 36, 132–146.

MacEwan, N. (2017). *Responsibilisation, rules and rule-following concerning cyber security: findings from small business case studies in the UK*. Doctoral dissertation. University of Southampton. Accessed 30.05.2021, https://eprints.soton.ac.uk/417156/1/Neil_MacEwan_Thesis_final_draft_post_viva_.pdf

Maimon, D., & Louderback, E. R. (2019). Cyber-dependent crimes: an interdisciplinary review. *Annual Review of Criminology*, 2, 191–216. https://doi.org/10.1146/annurev-criminol-032317-092057

Markopoulou, D., Papakonstantinou, V., & De Hert, P. (2019). The new EU cybersecurity framework: the NIS directive, ENISA's role and the General Data Protection Regulation. *Computer Law & Security Review*, 35(6). https://doi.org/10.1016/j.clsr.2019.06.007

McAlaney, J., & Hills, P. J. (2020). Understanding phishing email processing and perceived trustworthiness through eye tracking. *Frontiers in Psychology*, 11, 1756. doi: 10.3389/fpsyg.2020.01756

McGuire, M., & Dowling, S. (2013). *Cyber Crime: A Review of the Evidence (Summary of key Findings and Implications)*. Research Report, 75. London: Home Office.

MediaPro. (2020). *2020 State of Privacy and Security Awareness Report.* MediaPro & Osterman Research, Inc.

Miró-Llinares, F. (2015). That cyber routine, that cyber victimization: profiling victims of cybercrime. In R. G. Smith, R. C. C. Cheung, & L. Y. C. Lau (Eds), *Cybercrime Risks and Responses: Eastern and Western Perspectives* (pp. 47–63). Hampshire: Palgrave Macmillan.

Moseley, A. (2021). Nudging in public policy. *Oxford Research Encyclopedia of Politics*. Accessed 20.06.2021, https://doi.org/10.1093/acrefore/9780190228637.013.949

Naffine, N. (2003). Who are law's persons? From Chesire Cats to responsible subjects. *The Modern Law Review*, 66(3), 346–367. https://doi.org/10.1111/1468-2230.6603002

Nguyen, C., Jensen, M. L., Durcikova, A., & Wright, R. T. (2020). A comparison of features in a crowdsourced phishing warning system. *Information Systems Journal*, 31(3), 473–513. https://doi.org/10.1111/isj.12318

Norris, G., Brookes, A., & Dowell, D. (2019). The psychology of internet fraud victimisation: A systematic review. *Journal of Police and Criminal Psychology*, 34, 231–245. https://doi.org/10.1007/s11896-019-09334-5

Patel, P. (2017). Forced sterilization of women as discrimination. *Public Health Reviews*, 38. https://doi.org/10.1186/s40985-017-0060-9

Patel, K. K., & Reichardt, S. (2016). The dark side of transnationalism social engineering and Nazism, 1930s-40s. *Journal of Contemporary History*, 51(1), 3–21. https://doi.org/10.1177/0022009415607956

Pease, K. (2001). Crime futures and foresight: challenging criminal behaviour in the information age. In D. Wall (Ed), *Crime and the Internet* (pp. 18–28). London: Routledge.

Perkins, R. C., Howell, C. J., Dodge, C. E., Burruss, G. W., & Maimon, D. (2020). Malicious spam distribution: a routine activities approach. *Deviant Behavior*. https://doi.org/10.1080/01639625.2020.1794269

Posick, C. (2018). *The Development of Criminological Thought: Context, Theory and Policy*. Oxon: Routledge.

Pratt, T. C., & Turanovic, J. J. (2016). Lifestyle and routine activity theories revisited: the importance of 'risk' to the study of victimisation. *Victims & Offenders*, 11(3), 335–354.

Priezkalns, E. (2019). Europe should act on call ID spoofing now. Accessed 29.06.2021, https://commsrisk.com/europe-should-act-on-call-id-spoofing-now/

Proofpoint. (2019). *Human Factor Report 2019*. Proofpoint, Inc.

Proofpoint. (2021). The first step: initial access leads to ransomware. Accessed 29.06.2021, https://www.proofpoint.com/us/blog/threat-insight/first-step-initial-access-leads-ransomware

PurpleSec. (2021). 2021 Cyber Security Statistics: The Ultimate List of Stats, Data & Trends. Accessed 30.05.2021, https://purplesec.us/resources/cyber-security-statistics/

Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act) (7.06.2019). *Official Journal of the European Union*, L 151, pp. 15–69. Accessed 29.05.2021, http://data.europa.eu/eli/reg/2019/881/oj

Reynald, D. M. (2009). Guardianship in action: developing a new tool for measurement. *Crime Prevention and Community Safety*, 11(1), 1–20.

Rigotti, E., & Rocci, A. (2006). Towards a definition of communication context. Foundations of an interdisciplinary approach to communication. *Studies in Communication Sciences*, 6(2), 155–180.

RiskIQ. (2020). *Ransomware attacks the next consequence of the coronavirus outbreak*. RiskIQ I3 Intelligence Brief.

Roberts-Miller, P. (2019). *Rhetoric and Demagoguery*. Carbondale: Southern Illinois University Press.

Salisbury, J. (2020, April 3). Warning over coronavirus scam texts which demand money 'for leaving the house'. *Southwark News*. Accessed 29.05.2021, https://www.southwarknews.co.uk/news/warning-over-coronavirus-scam-texts-which-demand-money-for-leaving-the-house/

Sasse, A., & Smith, M. (2016). The security-usability tradeoff myth. *IEEE Security & Privacy*, 14(5), 11–13. doi: 10.1109/MSP.2016.102

Schmidt, A. T., & Engelen, B. (2020). The ethics of nudging: an overview. *Philosophy Compass*, 15(4). https://doi.org/10.1111/phc3.12658

Schreier, M. (2014). Qualitative content analysis. In U. Flick (Ed), *The SAGE Handbook of Qualitative Data Analysis* (pp. 170–183). Thousand Oaks, CA: SAGE Publications, Inc.

Sirks, B. (2013). Reform and legislation in the Roman Empire. *Mélanges de l'École française de Rome – Antiquité*, 125(2). https://doi.org/10.4000/mefra.1871

Smahel, D., MacHackova, H., Mascheroni, G., Dedkova, L., Staksrud, E., Olafsson, K., Livingstone, S., & Hasebrink, U. (2020). *EU Kids Online 2020: survey results from 19 countries*. London: London School of Economics and Political Science.

Solove, D. J. (2011). *Nothing to Hide: The False Tradeoff Between Privacy and Security*. New Haven: Yale University Press.

Sootak, J. (2010). *Karistusõigus. Üldosa.* [Criminal Law. General Part]. Tallinn: Juura.

Spano, R., & Freilich, J. D. (2009). An assessment of the empirical validity and conceptualization of individual level multivariate studies of lifestyle/routine activities theory published from 1995 to 2005. *Journal of Criminal Justice*, 37, 305–314. doi: 10.1016/j.jcrimjus.2009.04.011

Spano, R., & Nagy, S. (2005). Social guardianship and social isolation: an application and extension of lifestyle/routine activities theory to rural adolescents. *Rural Sociology*, 70, 414–437.

Steinmetz, K., Pimentel, A., & Goe, W. R. (2021). Performing social engineering: a qualitative study of information security deceptions. *Computers in Human Behavior*, 124, 106930. doi: https://doi.org/10.1016/j.chb.2021.106930

Stockman, M., Nedelec, J., & Mackey, W. (2016). Organizational cybervictimization: data breach prevention using a victimological approach. In T. J. Holt (Ed), *Cybercrime Through an Interdisciplinary Lens* (pp. 127–149). Oxon: Routledge.

Stojnic, T., Vatsalan, D., & Arachchilage. (2021). Phishing email strategies: understanding cybercriminals' strategies of crafting phishing emails. *Security and Privacy*, e165. https://doi.org/10.1002/spy2.165

Swinhoe, D. (2020). Pharming explained: how attackers use fake websites to steal data. Accessed 29.06.2021, https://www.csoonline.com/article/3537828/pharming-explained-how-attackers-use-fake-websites-to-steal-data.html

Turkle. (2011). *Alone Together: Why We Expect More from Technology and Less from Each Other*. New York: Basic Books.

Van Hoecke, M. (2011). Legal doctrine: which method(s) for what kind of discipline? In M. Van Hoecke (Ed), *Methodologies of Legal Research: Which Kind of Method for What Kind of Discipline?* (pp. 1–18). Oxford: Hart Publishing Ltd.

Verma, R., Crane, D., & Gnawalli, O. (2018). Phishing during and after disaster: Hurricane Harvey. *Resilience Week (RWS)*, 88–94. doi: 10.1109/RWEEK.2018.8473509

West, R. L, & Turner, L. H. (2019). *Introducing Communication Theory: Analysis and Application* (6th ed). New York, NY: McGraw-Hill Education.

Whitty, M. T. (2019). Predicting susceptibility to cyber-fraud victimhood. *Journal of Financial Crime*, 26(1), 277–292. doi: 10.1108/JFC-10-2017-0095

WHSV. (2020). Scammers take new approach to classic utility scam amid COVID-19. Accessed 29.05.2021, https://www.whsv.com/content/news/Scammers-take-new-approach-to-classic-utility-scam-amid-COVID-19-569985421.html

Wilcox, P, & Cullen, F. T. (2018). Situational opportunity theories of crime. *Annual Review of Criminology*, 1, 123–148. https://doi.org/10.1146/annurev-criminol-032317-092421

Williams, E. J., Hinds, J., & Joinson, A.N. (2018). Exploring susceptibility to phishing in the workplace. *International Journal of Human-Computer Studies*, 120, 1–13. https://doi.org/10.1016/j.ijhcs.2018.06.004

Williams, E. J., & Polage, D. (2019). How persuasive is phishing email? The role of authentic design, influence and current events in email judgements. *Behaviour & Information Technology*, 38(2), 184–197. https://doi.org/10.1080/0144929X.2018.1519599

Wright, R. T., Jensen, M. L., Thatcher, J. B., Dinger, M., & Marett, K. (2014). Research note – influence techniques in phishing attacks: an examination of vulnerability and resistance. Information Systems Research, 25(2), 385–400.

Yar, M. (2005). The novelty of 'cybercrime': an assessment in light of routine activity theory. *European Journal of Criminology*, 2(4), 407–427. doi: 10.1177/147737080556056

Zielinska, O. A., Welk, A. K., Mayhorn, C. B., & Murphy-Hill, E. (2016). A temporal analysis of persuasion principles in phishing emails. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 60(1), 765–769.

SUMMARY IN ESTONIAN

Kuritegevus kui kommunikatsioon: diagnostiliselt kasuliku teabe tuvastamine manipulatsioonirünnete sisust ja kontekstist

Tänapäevases teabe ülekülluse tingimustes on aina sagedasemaks muutunud erinevat tüüpi manipulatsiooniründed, mis jõuavad sõnumisaajateni e-kirjade, telefoni, lühisõnumite kui ka sotsiaalmeedia vahendusel. Tähelepanu hajutatuse, saabunud sõnumite tekitatud tugevate emotsioonide ning tõerääkimise eeldamise koostoime tõttu on manipulatsioonirünnete ohvriks langemine üha süvenev probleem, millega kaasneb nii majanduslik kui ka vaimne kahju. **Doktoritöö eesmärk oli teada saada, milline ründe ennetamise aspektist oluline teave on sõnumisaajatele kättesaadav manipulatsioonirünnete toimepanemiseks kasutatud sõnumite sisust ja kontekstist**.

Lähtudes eeldusest, mille kohaselt tuleneb ja sõltub iga tehnoloogia vahendatud keskkonnas toimuv tegevus kommunikatsioonist, olen doktoritöös esitanud **kuritegevus kui kommunikatsioon** (*crime-as-communication*) käsituse loomiseks vajaliku teoreetilise tausta, tehtud uurimustest saadud tulemused ning käsituse vajalikkust põhjendava arutelu. Kommunikatsioonina toime pandud kuritegevuse tuvastamisel ja tõrjumisel on võtmetähtsusega, et sõnumite vastuvõtja mõistaks veenmistaktikaid, millele süüteo toimepanijad tuginevad, et ajendada sõnumisaajaid tegema endale kahjulikku tegu, ning seda, kuidas kõnealused veenmistaktikad manipulatsioonirünnetes kasutatavates sõnumites esinevad.

Töö aluseks olevates uurimustes kogutud e-kirjade (**Uurimus I**) ja rahvusvahelistes meediaväljaannetes kajastatud pettuste kirjelduste (**Uurimus II** ja **III**) analüüsimiseks kasutasin kvalitatiivset tekstianalüüsi ning kvalitatiivset ja kvantitatiivset sisuanalüüsi. **Uurimus IV** on oma sisult teooriakriitika, mis on vajalik, et põhjendatult liikuda küberkuritegevuse olemasolevate käsituste juurest kuritegevus kui kommunikatsioon käsituseni. Püstitasin doktoritöös kaks peamist uurimisküsimust, millele andsin töö tulemusel alljärgnevad vastused:

1. Millist diagnostiliselt kasulikku teavet saavad sõnumisaajad manipulatsioonirünnete sisust ja kontekstist?

a. Kuidas aitavad karistusõiguses määratletud konventsionaalsed ehk tavapärased süüteokoosseisud mõista vahendatud kujul esinevat kuritegevust?

Kelmuse ja väljapressimise süüteod annavad olulise panuse kuritegevus kui kommunikatsioon käsituse mõistmiseks, kuna kõnealused süüteod põhinevad süüteo sihtmärgi (sõnumi vastuvõtja) teotahte mõjutamisel (**Uurimus I**) ning eeldavad sihtmärgilt teatud teo tegemist. Nimetatud eeldused on manipulatsioonirünnete määravad tunnused. Üldise suhtlusviisi, mida manipulatsioonirünnete toimepanijad kasutavad sõnumi vastuvõtjate teotahte mõjutamisel, ehk veenmise või ähvardamise kasutamise tuvastamist, tuleb pidada olulisemaks manipulatsioonirünnetes esinevatele konkreetsetele süžeedele keskendumisest.

b. Millised üldiseid kommunikatsioonivõtteid kasutatakse manipulatsioonirünnete eesmärgil koostatud sõnumites?

Pakkusin välja ja uurisin õngitsuskirjades esinevaid kelmus- ja väljapressimislaadseid kommunikatsiooni tüüpjuhtumeid (**Uurimus I**) ning kasutasin kõnealuseid tüüpjuhtumeid sisendina, et piiritleda ja selgitada kasusaamisel põhineva „hea samariitlase“ ja kahjukandmisega ähvardava „šokk ja ehmatus“ (ingl k *shock and awe*) sõnumi koostamise viiside kasutamist manipulatsioonirünnetes (**Uurimus II**). **Uurimus III** andis ülevaate „hea samariitlase“ ja „šokk ja ehmatus“ sõnumi koostamise viiside suhtelisest kasutamisest COVID-19 viiruspandeemia esimese nelja kuu jooksul, näidates, et 1040 manipulatsiooniründest 900 (86,5%) puhul lähtuti kasusaamisel põhinevast ning 140 (13,5%) puhul kahjukandmisega ähvardavast sõnumi koostamise viisist. Vähemalt ühe eespool kirjeldatud sõnumi koostamise viisi tuvastamine on manipulatsiooniründe tuvastamise mõttes diagnostiliselt kasulik teave.

c. Millist rolli omab manipulatsioonirünnetes matkimine?

Matkimine, sh sõnumi edastajat puudutava teabe võltsimine (ingl k *spoofing*), on kesksel kohal enamustes manipulatsioonirünnetes ning seda ka juhul, kui sõnumite saatjad otsustavad esineda väljamõeldud isikute või asutustena või pöörduvad sõnumi vastuvõtjate poole anonüümselt (**Uurimus I**). **Uurimused II** ja **III** näitasid, et sobivate ühiskondlike tingimuste esinemisel viivad küberkurjategijad manipulatsioonirünnete sisu, näiteks matkitavad isikud ja asutused, sageli sotsiaalse kontekstiga vastavusse.

d. Milliseid kanaleid kasutavad süüteo toimepanijad manipulatsioonirünnete toimepanemiseks?

Uurimus III näitas, et kuigi COVID-19 viiruspandeemia esimese nelja kuu jooksul olid e-kirjad populaarseim kanal manipulatsioonirünnete toimepanemiseks (53,5%), kasutati manipulatsioonirünnete toimepanemiseks ka telefonikõnesid (13,6%), lühisõnumeid (12,6%), võltsveebilehti (11,3%) ning sotsiaalmeedias edastatud sõnumeid ja postitusi (3,1%). Manipulatsioonirünneteks kasutatavate kanalite mitmekesisus annab põhjust konkreetse kanali asemel esmajärgus keskenduda sõnumi saatja ja vastuvõtja vahel olemasolevale avatud kanalile üldisemalt, sest avatud kanali kui sellise olemasolu aitab põhjendada, kui lihtsalt saavad küberkurjategijad oma sihtmärkidega tänapäeva infoühiskonnas manipulatsiooniründeks vajaliku ühenduse luua (**Uurimus IV**). Pettuse tuvastamiseks diagnostiliselt kasuliku teabe mõttes tuleks manipulatsioonründeid esmajärgus käsitada kanalineutraalsetena, st kanalite iseärasustest on esmajärgus olulisem rõhutada sõnumi saatja ja vastuvõtja vahelise ühenduse võimalikkust.

e. Millistele sisulistele teemadele tuginevad manipulatsioonirünnete toimepanijad oma sõnumites?

Uurimus I tulemused näitasid, et tavatingimustes kasutavad kurjategijad manipulatsioonirünnete sõnumites üldtuntud teemasid, muutes sõnumid sedasi asjakohasteks võimalikult suurele sõnumisaajate arvule, ning sarnaste teemade kasutamine võib manipulatsioonirünnete puhul esineda lainetena. **Uurimused II** ja **III** andsid kinnitust sellest, et konkreetsete ja manipulatsiooniründe mõttes asjakohaste ühiskondlike tingimuste olemasolul on küberkurjategijad suutelised sõnumites kasutatavad teemad ja sõnumite sisu kiiresti sotsiaalsele kontekstile sobivaks kohandama. Näiteks, COVID-19 viiruspandeemia esimestel kuudel usaldusväärset teavet otsivatele inimestele pakkusid kelmid „tervisenippe" ja pandeemia süvenemisel rahalisi toetusi. Samuti polnud kelmidel hetkekski tarneprobleeme kaitsemaskide, COVID-19 ravimite ning isegi vaktsiinidega. Eeltoodust järeldub, et kurjategijad peavad kõnealuseid kohandamisi kuritegelikul teel saadud kasu saamise mõttes kasumlikumaks kui seda on tavatingimustes kasutatavad üldised teemad.

f. Milliseid konkreetseid sotsiaalpsühhooloogilisi mõjustamistehnikaid manipulatsioonirünnetes kasutatakse?

Uurimuses III kasutasin kvalitatiivse sisuanalüüsi täiendava sammuna Cialdini kuut mõjustamispõhimõtet ning saadud tulemused andsid empiirilise kinnituse, et manipulatsioonirünnetes leidsid kasutust kõik kuus mõjustamispõhimõtet: autoriteetsus, nappus, meeldivus, sotsiaalne kinnitus, järjekindlus ning vastastikkus. Asjaolu, et empiirilist kinnitust leidis kõigi kuue mõjustamispõhimõtte kasutamine, rõhutab veelgi küberkurjategijate tegevuse laiahaardelisust manipulatsioonirünnetes kasutatavate sõnumite vastuvõtjate mõjutamisel. Teadmised Cialdini mõjustamispõhimõtete kasutamisest muudes valdkondades, näiteks kaubanduses, võivad anda praktilise viisi, kuidas koolitada inimesi samu mõjustamistehnikaid märkama ka manipulatsioonirünnete puhul.

g. Millist rolli omab manipulatsioonirünnetes sotsiaalne kontekst?

Uurimus I tulemused näitavad, et küberkurjategijad panustavad väljamõeldud lugudesse enam siis, kui puudub asjakohane sotsiaalne kontekst, mida loost tõepärase mulje jätmiseks ära on võimalik kasutada. Eeltoodud juhtumite puhul luuakse toimuva suhtluse kontekst saadetavate sõnumitega. **Uurimuste II** ja **III** tulemuste kohaselt ajendavad manipulatsioonirünnete mõttes asjakohases sotsiaalses kontekstis toimuvad (järsud) muutused, näiteks COVID-19 viiruspandeemia tekkimine, küberkurjategijaid manipulatsioonirünnete toimepanemiseks kasutatavate sõnumite sisu kõnealuste muutuste kajastamiseks kohandama. Taolised muudatused kajastusid sõnumites sageli kasutatud teemades, isikutes või asutustes, keda matkiti, ning kommunikatsioonivõtetes, millele sõnumite koostamisel tugineti.

2. Kuidas panustab kuritegevus kui kommunikatsioon käsitus praegusesse kriminoloogilisse mõtlemisse?

Selle asemel, et allutada küberkuritegevuse juhtumid nn füüsilisele (ehk konventsionaalsele) arusaamale tegevusest ja isikutevahelisest kokkupuutest, loob kuritegevus kui kommunikatsioon käsitus kommunikatsioonipõhise aluse tehnoloogia vahendatud keskkondades toimuvate tegevuste mõistmiseks. Küberkuritegelike tegevuste sisuks olevate kommunikatsiooniprotsesside esiletoomisega annab kuritegevus kui kommunikatsioon käsitus üksikasjaliku ja põhimõtetest lähtuva viisi nii tehnoloogia vahendatud kuritegevuse kui ka muude tegevuste mõtestamiseks ja analüüsimiseks. Eeltoodule lisaks rõhutab kuritegevus kui kommunikatsioon käsitus avatud kanalite ning sõnumite tõlgendamise olulisust küberkuritegevuse ohvriks langemise vältimises. Avatud kanali olemasolul on sõnumi vastuvõtja omadused – nii inimestel kui ka masinatel – kuritegevus kui kommunikatsioon käsituses olulised üksnes ulatuses, milles need mõjutavad saabuvate sõnumite tõlgendamist ehk kuritegeliku sõnumi või sisendi eristamist kogu saabunud teabest.

PUBLICATIONS

CURRICULUM VITAE

Name:	Kristjan Kikerpill
Date of birth:	November 5, 1989
E-mail:	kristjan.kikerpill@gmail.com

Education:

10.02.2020–	University of Tartu, PhD studies (sociology)
2013–2016	University of Tartu, MA studies in Law (information technology law)
2009–2013	University of Tartu, BA studies in Law

Fields of research:
cyber-criminology, sociology, communication

Publications:

Kikerpill, K. (2022). Misinformation, Scandalization, and the Trump Show: Audience Responses to President Trump's Pandemic Narrative. In: Sarina Chen, Zhuojun Joyce Chen, Nicole Allaire (Eds.), *Discordant Pandemic Narratives in the U.S*. Lexington Publishing [Rowman & Littlefield] (upcoming).

Kikerpill, K. & Siibak, A. (2021). Mazephishing: The COVID-19 pandemic as credible social context for social engineering attacks. *Trames Journal of the Humanities and Social Sciences* (upcoming).

Kikerpill, K., Siibak, A. & Valli, S. (2021). Dealing with Deepfakes: Reddit, Online Content Moderation, and Situational Crime Prevention. In: Julie B. Wiest (Ed.), *Theorizing Criminality and Policing in the Digital Media Age* (pp. 25–45). Emerald Publishing Limited. https://doi.org/10.1108/S2050-206020210000020008

Kikerpill, K. & Siibak, A. (2021). Abusing the COVID-19 Pan(dem)ic: A Perfect Storm for Online Scams. In: Pollock, J. C., Kovach, D. (Eds.), *COVID-19 in International Media: Global Pandemic Perspectives* (pp. 249–258). New York: Routledge. DOI: 10.4324/9781003181705-25

Kikerpill, K. (2021). The Individual's Role in Cybercrime Prevention: Internal Spheres of Protection and Our Ability to Safeguard Them. *Kybernetes*, 50(4): 1015–1026. DOI: 10.1108/K-06-2020-0335

Kikerpill, K. (2020). Choose Your Stars and Studs: The Rise of Deepfake Designer Porn. *Porn Studies*, 7 (4), 352−356. DOI: 10.1080/23268743.2020.1765851

Kikerpill, K. & Siibak, A. (2019). Living in a Spamster's Paradise: Deceit and Threats in Phishing Emails. *Masaryk University Journal of Law and Technology*, *13*(1), 45−63. DOI: 10.5817/MUJLT2019-1-3.

Kikerpill, K. (2019). Work, Prey, Love: A Critical Analysis of Estonian Cybercrime Case Law 2014–2019. *Proceedings: Estonian Academy of Security Sciences*, *18*, 109−138.

Mäses, S., Kikerpill, K., Jüristo, K. & Maennel, O. (2019). Mixed methods research approach and experimental procedure for measuring human factors in cybersecurity using phishing simulations. In: A. Stacey (Ed.), *Proceedings of the 18th European Conference on Research Methodology for Business and Management Studies, ECRM 2019* (pp. 218–226). Reading, UK: Academic Conferences and Publishing International Limited.

Teaching experience:

Lecturer: Overview of Public Communication and Media Regulation (Spring 2021); Information Law, Privacy and Data Protection (Fall 2021)

ELULOOKIRJELDUS

Nimi:	Kristjan Kikerpill
Sünniaeg:	5. november 1989
E-post:	kristjan.kikerpill@gmail.com

Haridustee:

10.02.2020–	Tartu Ülikool, doktoriõpe (sotsioloogia)
2013–2016	Tartu Ülikool, õigusteaduse magistriõpe (infotehnoloogiaõigus)
2009–2013	Tartu Ülikool, õigusteaduse bakalaureuseõpe

Peamised uurimisvaldkonnad:
küberkriminoloogia, sotsioloogia, kommunikatsioon

Publikatsioonid:

Kikerpill, K. (2022). Misinformation, Scandalization, and the Trump Show: Audience Responses to President Trump's Pandemic Narrative. In: Sarina Chen, Zhuojun Joyce Chen, Nicole Allaire (Eds.), *Discordant Pandemic Narratives in the U.S*. Lexington Publishing [Rowman & Littlefield] (ilmumas).

Kikerpill, K. & Siibak, A. (2021). Mazephishing: The COVID-19 pandemic as credible social context for social engineering attacks. *Trames Journal of the Humanities and Social Sciences* (ilmumas).

Kikerpill, K., Siibak, A. & Valli, S. (2021). Dealing with Deepfakes: Reddit, Online Content Moderation, and Situational Crime Prevention. In: Julie B. Wiest (Ed.), *Theorizing Criminality and Policing in the Digital Media Age* (lk 25–45). Emerald Publishing Limited. https://doi.org/10.1108/S2050-206020210000020008

Kikerpill, K. & Siibak, A. (2021). Abusing the COVID-19 Pan(dem)ic: A Perfect Storm for Online Scams. In: Pollock, J. C., Kovach, D. (Eds.), *COVID-19 in International Media: Global Pandemic Perspectives* (lk 249–258). New York: Routledge. DOI: 10.4324/9781003181705-25

Kikerpill, K. (2021). The Individual's Role in Cybercrime Prevention: Internal Spheres of Protection and Our Ability to Safeguard Them. *Kybernetes*, 50(4): 1015–1026. DOI: 10.1108/K-06-2020-0335

Kikerpill, K. (2020). Choose Your Stars and Studs: The Rise of Deepfake Designer Porn. *Porn Studies*, 7 (4), 352−356. DOI: 10.1080/23268743.2020. 1765851

Kikerpill, K. & Siibak, A. (2019). Living in a Spamster's Paradise: Deceit and Threats in Phishing Emails. *Masaryk University Journal of Law and Technology*, *13*(1), 45−63. DOI: 10.5817/MUJLT2019-1-3.

Kikerpill, K. (2019). Work, Prey, Love: A Critical Analysis of Estonian Cybercrime Case Law 2014–2019. *Proceedings: Estonian Academy of Security Sciences*, *18*, 109−138.

Mäses, S., Kikerpill, K., Jüristo, K. & Maennel, O. (2019). Mixed methods research approach and experimental procedure for measuring human factors in cybersecurity using phishing simulations. In: A. Stacey (Ed.), *Proceedings of the 18th European Conference on Research Methodology for Business and Management Studies, ECRM 2019* (lk 218–226). Reading, UK: Academic Conferences and Publishing International Limited.

Õpetamiskogemus:

Lektor: Infoõiguse alused (kevad 2021); Infoõigus, privaatsus ja andmekaitse (sügis 2021)

DISSERTATIONES SOCIOLOGICAE UNIVERSITATIS TARTUENSIS

1. **Veronika Kalmus.** School textbooks in the field of socialisation. Tartu, 2003, 206 p.
2. **Kairi Kõlves.** Estonians' and Russian minority's suicides and suicide risk factors: studies on aggregate and individual level. Tartu, 2004, 111 p.
3. **Kairi Kasearu.** Structural changes or individual preferences? A study of unmarried cohabitation in Estonia. Tartu, 2010, 126 p.
4. **Avo Trumm.** Poverty in the context of societal transitions in Estonia. Tartu, 2011, 215 p.
5. **Kadri Koreinik.** Language ideologies in the contemporary Estonian public discourse: With a focus on South Estonian. Tartu, 2011, 128 p.
6. **Marre Karu.** Fathers and parental leave: slow steps towards dual earner/ dual carer family model in Estonia. Tartu, 2011, 125 p.
7. **Algi Samm.** The relationship between perceived poor family communication and suicidal ideation among adolescents in Estonia. Tartu, 2012, 121 p.
8. **Tatjana Kiilo.** Promoting teachers' efficacy through social constructivist language learning: challenges of accommodating structure and agency. The case of Russian-speaking teachers in Estonia. Tartu, 2013, 156 p.
9. **Ave Roots.** Occupational and income mobility during post-socialist transformation of 1991–2004 in Estonia. Tartu, 2013, 130 p.
10. **Tarmo Strenze.** Intelligence and socioeconomic success A study of correlations, causes and consequences. Tartu, 2015, 119 p.
11. **Mervi Raudsaar.** Developments of social entrepreneurship in Estonia. Tartu, 2016, 141 p.
12. **Ero Liivik.** Otsedemokraatia Eestis: õigussotsioloogilisi aspekte. Tartu, 2017, 166 p.
13. **Mai Beilmann.** Social Capital and Individualism – Collectivism at the Individual Level. Tartu, 2017, 145 p.
14. **Rainer Reile.** Self-rated health: assessment, social variance and association with mortality. Tartu, 2017, 123 p.
15. **Katri Lamesoo.** Social Construction of Sexual Harassment in the Post-Soviet Context on the Example of Estonian Nurses. Tartu, 2017, 185 p.
16. **Andu Rämmer.** Sotsiaalse tunnetuse muutused Eesti siirdeühiskonna kontekstis. Tartu, 2017, 230 p.
17. **Kadri Rootalu.** Antecedents and consequences of divorce in Estonia from longitudinal and multigenerational perspectives. Tartu, 2017, 128 p.
18. **Kairi Talves.** The dynamics of gender representations in the context of Estonian social transformations. Tartu, 2018, 129 p.
19. **Aare Kasemets.** Institutionalisation of Knowledge-Based Policy Design and Better Regulation Principles in Estonian Draft Legislation. Tartu, 2018, 252 p.

20. **Dagmar Narusson**. Personal-recovery and agency-enhancing client work in the field of mental health and social rehabilitation: Perspectives of persons with lived experience and specialists. Tartu, 2019, 139 p.
21. **Oliver Nahkur.** Measurement of Interpersonal Destructiveness: the Societal Perspective. Tartu, 2019, 164 p.
22. **Tayfun Kasapoglu**. Algorithmic Imaginaries of Syrian Refugees: Exploring Hierarchical Data Relations from the Perspective of Refugees. Tartu, 2021, 152 p.