AALBORG UNIVERSITY
DENMARK

Semester:
10th Semester

Title:
An Analysis in Current Social Engineering Attacks and Countermeasures

Project Period:
Spring Semester 2023

Semester Theme:
Thesis Project

Participent
Hac Memet Duvarci,
No. 20176600

Supervisor:
Marios Anagnostopoulos

Copies:
1

Page Number: 79 pages

Date of Completion:
October 18th, 2023

Abstract

This report addresses the multifaceted domain of social engineering attacks and the countermeasures employed to mitigate them, focusing on both technical and non-technical strategies. Firstly, the study is informed by a blend of methodological approaches, including expert interviews and a survey targeting professionals. The data generated serves as a cornerstone for identifying common types of social engineering attacks, challenges in implementing countermeasures, and the effectiveness of existing strategies. The report's outcome is a set of recommendations aimed at enhancing an organization's resilience against social engineering attacks. These include best practices for integrating technical and non-technical countermeasures into an existing cyber security strategy.

An Analysis in Current Social Engineering Attacks and Countermeasures

Hac

Aalborg University Copenhagen

Contents

Chapter 1

Introduction

As the digital landscape continues to mature and organizations increasingly rely on technology to support their business operations, the risk of cyber attacks has become a critical concern. Social engineering attacks have emerged as an effective form of cyber attack, increasing in intensity and number, leveraging the psychological vulnerabilities of human vulnerabuilities to gain unauthorized access to information or systems. These attacks can be sophisticated, difficult to detect, and can have significant financial, reputational, and legal consequences for organizations.

Organizations must develop and implement effective countermeasures to combat social engineering attacks that can prevent, detect, and respond to attacks. However, the efficacy of these countermeasures can be challenging to determine as social engineering tactics continue to evolve and attackers become increasingly sophisticated in their approach. Additionally, organizations may need help implementing and maintaining countermeasures, including limits in resources, lack of cybersecurity expertise, and employee awareness and training. The problem is identifying the most effective technical and non-technical countermeasures to defend against social engineering attacks and examining the challenges organizations face in implementing and maintaining these measures. Here, the project aims to explore the various types of social engineering attacks and the technical countermeasures used to defend against them. In addition to technical countermeasures, non-technical countermeasures such as employee training, awareness programs, and policies and procedures will also be examined in this study. Furthermore, the project will explore the role of employee awareness and training in preventing successful social engineering attacks and the challenges in delivering social engineering-preventing software.

The research aims to provide insights and best practices for organizations to de-

velop and implement effective countermeasures to defend against social engineering attacks. This study's findings will potentially benefit organizations, policymakers, and cybersecurity professionals to associate various social engineering attacks in the cybersecurity spectrum, contributing to the understanding of social engineering attacks and their associated countermeasures. By outlining the various practice and recommendations for integrating countermeasures into an organization's cybersecurity strategy, this study will help organizations better protect themselves against social engineering attacks and improve their overall cybersecurity posture.

1.1 Problem formulation

How do organizations approach raising cyber awareness and strengthening resilience against social engineering attacks?

1. What are the most common social engineering techniques used by attackers, and how do they influence user behaviour and decision-making in organizations?
2. What are the most common countermeasures (ML, user-training) used to defend against social engineering attacks, and how effective are they?
3. What are the best practices for organizations in raising awareness about social engineering and improving their overall cyber hygiene?

1.1.1 Contribution

This research project addresses a pressing concern in the field of cyber security: social engineering attacks and the countermeasures that can be used to prevent them. It bridges a critical gap in the existing literature by associating various social engineering attacks within the broader cyber security spectrum. The approach to analysing technical and non-technical countermeasures offers a multi-dimensional perspective on the defence mechanisms. Furthermore, the insights drawn from the surveys and interviews are valuable. By capturing real-world opinions and experiences, this research offers a pragmatic view of organisations' challenges. This ensures that the recommendations made are theoretically sound and grounded in the reality of the cybersecurity industry.
Moreover, the outlined recommendations for integrating countermeasures into an organisation's cyber security strategy aim for organisations seeking to enhance their security posture against social engineering attacks. The research project contributes to the field of cyber security by offering a current understanding of social engineering attacks, outlining effective countermeasures, and providing pragmatic

recommendations to strengthen an organisation's defence mechanisms. Through these contributions, the study aims to assist those who find themselves impacted or influenced by social engineering attacks, offering guidance as they navigate the complexities of these threats and work towards a safer digital environment.

1.1.2 Limitations

The scope and depth of this study were influenced by a set of limitations that need to be acknowledged. The main challenge was the lack of cooperation from companies, which limited access to diverse insights and understanding of social engineering defence mechanisms. Such reluctance likely stems from concerns about revealing sensitive data. For the investigation of social engineering countermeasures, despite a thorough approach, there was a lack of depth in the search for available countermeasures as well as the countermeasures presented. The landscape of social engineering attacks suggests that the numerous techniques and countermeasures are likely not addressed within this study, marking a potential area for future research. The insights gained from the conducted survey and interviews for this report provided, but acquiring such diverse data sources also highlighted another limitation which was the need for a more in-depth investigation. The insights gained could have been bolstered by further exploring the complexities of the interviews and survey responses, especially in understanding how various perspectives align or diverge. This deeper exploration could potentially offer a richer understanding of how to enhance social engineering countermeasures.

While this study offers a broad glance at social engineering attacks and their countermeasures, it acknowledges the inherent limitations imposed by the available resources, access to diverse company profiles, and the depth of investigation carried out.

1.2 Structure of the report

For the purpose of establishing a common understanding of the key terms and concepts related to social engineering and countermeasures, we first explore some basal knowledge on the topic.

The following chapter, chapter 2, is intended to provide a strong foundational understanding of social engineering attacks and countermeasures. It is presented by defining key terms and concepts related to social engineering. The chapter also explores common technical as well as non-technical countermeasures used to

defend against social engineering attacks and the associated countermeasures. Finally, the chapter intends to provide a solid understanding of social engineering attacks. By providing this foundational understanding, the report provides a better grasp of the subject matter before approaching Chapter 3, the literature review, and Chapter 4, methodology, later in the report.

Chapter 3 reviews the literature on social engineering attacks and their countermeasures. Opposed to Chapter 2, this chapter exclusively relies on academic papers. It covers definitions, types, and methods of social engineering attacks, as well as the challenges organizations face in defending against them. The chapter also reviews the state of the art in social engineering countermeasures, including technical and non-technical approaches.

Chapter 4 will describe the research methodology used in the project. It introduces the research approach and elaborates on the chosen research methods and design, analysis methods, as well as the data collection process described in iterations.

Chapter 5 intends to provide the rationale for the chosen collection process. Here, the interview questions are explained as well as the survey questions.

Chapter 6 contains discussions of the data collection findings which are used as a baseline for the conclusion. This chapter also includes the project's findings.

Chapter 7 contains the conclusion and discussion of the project.

Chapter 2

What is Social Engineering

2.1 Introduction to social engineering

While "hacking" generally refers to a range of techniques targeting computer systems, social engineering distinctly leverages human interaction to achieve unauthorized access, whether for sensitive information, financial benefits, or political gains[1]. Social engineering has elements from both psychology and information security. One definition states that "Social Engineering uses influence and persuasion to deceive people by convincing them that the social engineer is someone he is not, or by manipulation"[2], while another describes it as "Social engineering is an attack on information security for accessing systems or networks"[3]. These quotes share a common concept: social engineering entails the ability to manipulate a target for a specific purpose.

Social engineering can be applied wherever there is a potential benefit for the social engineer. However, it is essential to note that social engineering is not inherently malicious; it is utilized in everyday interactions within families, by doctors, or in non-malicious transactions. In these cases, the social engineer leverages the dynamics of their relationship with the target to achieve a particular outcome. Examples include parents guilt-tripping their children, or a customer service representative maintaining a calm demeanour to encourage customers to mirror their behaviour or a news broadcast cherry-picking information to capture viewers' attention [4].

The Common Vulnerability Scoring System (CVSS), a scoring system based on vulnerability severity, assigns social engineering attacks as an attack vector, stating that the attacker relies on user interaction to manipulate a target at a metric value classification[5]. Thus, rather than using computers as the main attack vector for a

cyber attack, social engineering is the means of using humans as the attack vector.

2.1.1 Importance of Studying Social Engineering

Social engineering's prominence in cyber attacks cannot be understated. An astonishing 98% of all cyber attacks are believed to incorporate some form of social engineering, as reported by security vendor PurpleSec[6]. Although a social engineering attack may not directly harm a target by itself, it has been demonstrated to play a significant role in major cyber attacks. The extent of social engineering's involvement varies, but its presence is unquestionable. Alternative estimates suggest that its influence on cyber attacks ranges from 70%, according to Proof-Point[7], to as high as 70-90% according to KnowBe4 [8]. These figures align with broader concerns in the cyber security community, as evidenced by the World Economic Forum (WEF) Global Cyber Security Outlook 2022 report. The report identifies social engineering as the second biggest concern among cyber leaders, just behind ransomware, followed by malicious insider activities. In contrast, the same survey reveals that 50% of all respondents find responding and recovering from a cyber incident challenging, and fewer than 25% of respondents from companies with 5,000 to 50,000 employees have the necessary people and skills in place today[9].

2.2 Types of social engineering attacks

Social engineering attacks can be distinguished by their platform and execution methods. The term social engineering covers a variety of tactics, but all social engineering attacks generally involve deceiving victims into performing an undesired action based on the attacker's chosen delivery method[10]. Although social engineering is typically conducted through emails, it can also occur via messages and telecommunication as a means of employing deceptive methods. As technology advances, computers and mobile devices are increasingly becoming interconnected, allowing the use of various applications and software across different platforms. This development makes social engineering attacks more convenient and potentially more effective, as attackers can exploit the seamless integration of these technologies to their advantage[11].

The following section will explore various social engineering attacks and examples of how they can be conducted. Given the multifaceted and evolving nature of social engineering, many attacks often involve multiple phases, integrating several tactics within a single operation. Here, the categories are platforms and execution methods, respectively. Platforms represent the distribution method of an attack,

while the execution method entails various tactics. The categorisation presented here aims to provide a structured overview. However, it is crucial to recognise that this is just one perspective. Different literature or resources might classify these attacks differently.

2.3 Platforms

Email phishing

Email phishing attacks are the most prevalent type and have become increasingly common. In Denmark, over half of the population received phishing emails between 2019 and 2020[12]. Business email compromise (BEC), a specific kind of phishing attack targeting organizations, is also widespread. According to Proofpoint, 83% of surveyed organizations experienced a successful phishing email attack, a substantial increase from 57% the previous year[10]. Among these victims, 11% faced ten or more successful attacks. Although BEC attacks are closely monitored and filtered, a small number of successful attempts can result in significant damage. Due to the electronic nature of emails, it is relatively easy to incorporate attachments and links with malware. One report states that most cyber attacks start with phishing mail[13], while another states phishing emails remain the primary method for ransomware delivery[14].

Vishing

Vishing, or Voice Phishing, is where a social engineer uses a cell phone or other similar means to deceive the victim. An example of a case is where the attacker used "deep voice" technology to clone the speech of a company director and trick a bank manager into thinking he was speaking with the actual director. The attacker requested a transfer of $35 million as part of a supposed acquisition by his company and sent emails confirming the details, which appeared legitimate to the bank manager. The bank manager made the transfers, unaware that he was being tricked. The attackers' use of advanced technology made it difficult for the bank manager to recognize the scam. [15] Vishing is still a relevant form of social engineering, and of the survey organizations on a global average, 75% are not aware of what vishing is[10].

Smishing

Smishing, also referred to as SMiShing (a reference to SMS), is a phishing attack transmitted through text messages[10]. Typically, the goal of this attack is to get

the target to interact with the content provided by the social engineer, whether by responding to the text message or clicking on a link that redirects to a malicious website, potentially providing the attacker with login credentials. Social engineers also take advantage of national events as they happen. Among many cases, one example is the NemID to MitID transition as evident in Fig. 2.1[16].

Figure 2.1: Smishing attack using the MitID transition as an attack vector [16]

2.4 Execution Methodology

2.4.1 Spear-phishing

Spear-phishing is a targeted form of phishing in which attackers research their victims to create specialised attacks that appear more personalised and believable. These attackers can use information from various sources gained legally through public information or otherwise illegally and aim to target specific individuals within an organisation[17]. Other attacks associated with spear-phishing are whaling and BEC, which are more specialised attacks. BEC attackers are considered to impersonate well-known brands or services, and whaling is considered to be targeting high-profile individuals[10][18]. It is clear that spear-phishing often includes the element of impersonating trusted entities, aiming at specific individuals within an organisation.

2.4.2 Impersonation

Impersonation in social engineering is the act of masquerading as another person, usually someone with authority or trust, to manipulate individuals into divulging information or performing actions. For example, in 2020, attackers gained unauthorized access to high-profile Twitter accounts, including those of Joe Biden and Barack Obama, and used these personas to request cryptocurrency transactions[19]. While the primary objective of this particular attack was financial gain, it could also serve other purposes, such as generating propaganda or conducting espionage (Fig. 2.2). Impersonation is not limited to individuals; in fact, a substantial portion of spear phishing attacks, constituting no less than 40%, impersonate larger organizations. Among these, Microsoft emerged as the most frequently impersonated company[18]. These attacks leverage the trust and authority associated with such organizations to deceive and exploit victims.

Figure 2.2: Joe Biden's and Barack Obama's Twitter accounts asking for Bitcoin [19]

2.4.3 Baiting

Baiting exploits a victim's curiosity or greed, tempting them to interact with a malicious objective. In some instances, a social engineer might leave a malware-infected USB drive in a particular place, hoping a curious individual will insert it into their computer[20]. This is usually referred to as USB baiting. In 2016, Google researchers conducted an experiment where they scattered 300 USB drives around a university campus, discovering that almost half were picked up and plugged into computers within minutes[21]. However, this technique extends beyond physical

devices and can also include online activities, such as clickbait or malicious links leading to harmful websites. The primary goal of baiting is to tempt the victim into engaging with the malicious item, thereby compromising their security. According to a survey conducted by Proofpoint, which involved organizations from seven EU countries, the UK, and the USA, 64% of the organizations experienced USB-based baiting attacks, which is a 15% increase from 2020 to 2021[10].

2.4.4 Bulk-phising

Bulk-phishing is a type of attack in which a large number of non-specific targets receive deceptive messages, often sent indiscriminately. This approach can be employed across various platforms, such as emails, text messages, or social media messages. A well-known example of bulk-phishing is the infamous Nigerian prince scam. According to Proofpoint, 86% of organizations have faced bulk phishing attacks[10].

Summary

Various social engineering are outlined and categorized by platform and execution methods. While this is not an exhaustive list, it gives an overview of social engineers' primary tactics, setting the stage for subsequent sections on mapping these attacks within broader cyber security contexts.

2.5 Mapping Social Engineering in Cyber Secrity

2.5.1 Cyber Kill Chain

The Cyber Kill Chain, introduced by Lockheed Martin in 2011, represents a sequence of stages of a cyberattack.citeCrowdstrikeCBC. By breaking down the various stages of a cyber attack, whether the attack is malware-based or an undetected network attack, organisations may improve their ability to intervene or halt potential threats[22]. See Fig below.

Figure 2.3: Stages of Cyber Kill Chain

The framework is initially intended for Advanced Persistent Threats (APT) deploying sophisticated malicious techniques. However, at certain stages of a cyberattack, attackers might lean towards social engineering tactics, recognising that manipulating the human nature may be more convenient than breaching technical defences. While each stage of the Cyber Kill Chain has its own importance, social engineering mostly stakes its claim in specific stages, notably the "Reconnaissance" and "Delivery" stages[23, 24].

In the reconnaissance stage of the Cyber Kill Chain, also called the research stage, attackers identify potential targets and search for vulnerabilities or details that might facilitate a breach[23]. Although there are technical tools designed for this purpose, such as Nmap or Wireshark, attackers may also use non-technical methods[24]. For instance, attackers might browse social media platforms to search for personal information and credentials or gather insights about software applications and operating system details or employ tactics such as phishing and smishing.

In the delivery stage, the appeal of using sophisticated computational attacks may seem essential, yet, a simple operation will do the trick. It is the simplicity of social engineering techniques, like sending a single deceptive email, that can prove most effective. Phishing and similar tactics emerge as pivotal tools for attackers in this stage, which highlights the imperative for organisations to not only recognise but also proactively counter these seemingly straightforward but potentially devastating scenarios[23].

However, while the reconnaissance stage may not predominantly rely on social engineering techniques, it's worth noting that such tactics can partly contribute to laying the groundwork for sophisticated attacks that manifest during the delivery

stage.

2.6 The impact of social engineering attacks on organizations

Organizations undoubtedly strive to avoid financial losses, reputational damage, and the costs associated with recovering from cyber incidents. In this context, social engineering is considered a significant cyber threat that demands attention. The average cost of a successful social engineering attack is around $130,000, although for larger businesses, this figure can easily reach millions or even billions of dollars.[**empty citation**] Naturally, companies invests in the prevention and mitigation form in any form of cyber attacks to protect their valuable assets and maintain a secure environment. These measures can be considered as an impact of the organizations, whether it is in the form of cyber awareness training, cyber insurance or in the rules of work policies. Though these measure may act not as intended, e.g. according to a survey by Barracuda, organizations which have drawn a cyber insurance is most likely to be victim of an ransomware attack including in paying the ransom to the attackers, which could mean that these organizations are selective to cyber criminals.

2.7 Social Engineering Countermeasures

There are several countermeasures that companies use to bolster their defence. Broadly, these can be categorized into two types: non-technical measures, which focus on the human aspect, and technical measures related to computer systems and technologies. Below is a brief overview of the types of countermeasures. It's worth noting that the countermeasures listed are not in-depth analysis of what is implemented in practice. In real-world scenarios, both types of measures may be used in combination as a defence strategy.

2.8 Non-technical countermeasures against social engineering attacks

While most people think cybersecurity is about advanced firewalls and updated software, the human factor should not be ignored. Often called the 'weakest link,'

individuals add an unpredictable element to any security setup. Being unpredictable opens up vulnerabilities, often making humans an easier target for hackers than the technology they use.[25]

According to Fortinet, for organisations which had experienced a cyberattack in the past 12 months, 81% of the attacks took the form of phishing attacks, password attacks and malware attacks[26]. These types of attacks include exploiting the weaknesses of employees. Employees stand at the forefront, bridging the gap between potential threats and an organization's internal systems. Thus, being proactive and enhancing the employees' ability to conduct their work in a safe manner is essential, particularly as the complexity and subtleties of social engineering strategies continue to change[26].

In the following sections, a range of social engineering countermeasures and principles for fortifying against social engineering will be explored, from awareness training to the establishment of policies and procedures.

2.8.1 Cyber Security Awareness Training

Cyber Awareness training, also called security awareness training, is designed to equip individuals with the right knowledge and intuition to recognize cyber threats in any form. Such training is not exclusive to social engineering but can be included as part of the overall awareness of cyber threats. Initially, organizations must map out their specific threat landscape, which involves evaluating current and anticipated challenges unique to their operations.[25] This means assessing the various communication channels and technologies in use, and understanding the corresponding potential threats. For instance, if an organisation leans heavily on mobile-based communications, it should remain vigilant against vishing and smishing. Similarly, if they are predominantly email-centric, the threat of email phishing ought to be a pronounced concern. So, the key is understanding the organization's assets and potential vulnerabilities, including the types of attacks employees can face.[10].

Traditional cyber training, i.e. industry cyber certification, alone may not be enough to adequately prepare organizations for the evolving threats they face, which have been found to often fail in translating to practical expertise, causing a gap between how prepared they seem and how skilled they actually are.[27] At bare minimum, employees needs to understand social engineering attacks are, to understanding the consequence of an adversary knowing the name of their dumpster vendor.[28]

Social engineering training includes responding to manipulative tactics that

exploit human psychology which emphasizes attackers' various techniques, from phishing emails to pretexting, enabling employees to distinguish between legitimate requests and deceptive tactics.

2.8.2 Test Scenario Simulations

The effectiveness of awareness training is often influenced by human factors, such as an individual's interest level, memory or concentration span, attributes that social engineers may exploit[29]. Organizations can incorporate scenario simulations into their training programs to bridge these challenges[10].

These simulations mimic real-world cyber threats, offering hands-on experience. Engaging employees in such realistic setups gives them a more profound awareness of potential attacks. Additionally, to make employees more invested and interested, organizations use material which engages emotions and triggers motivation to be aware of decision-making[25].

The success of these simulation programs is naturally measured by the change in employee behavior during real incidents. However, the tangible results of such training might only become evident in the long term. Moreover, regularly updating and conducting these tests is important. This ensures that employees always remain alert and aware of the evolving threat landscape[10].

2.8.3 (Cyber) Security Policies

Cybersecurity policies encapsulate a set of guidelines, rules, and restrictions that protect both employees and the organization. These policies are designed to help employees make decisions in social engineering-related cases[30]. In other words, when the employee needs to make a decision, policies take the thinking out of the equation[28].

These policies can be expressed through various measures, such as policies about reporting, mobile device management, multi-factor authentication, password management, and keeping software up-to-date. While some of these measures rely on technology, it is ultimately the individuals behind the technology who must act cohesively and follow these guidelines and restrictions. In the era of remote work, where companies may rely on digital communication methods such as email, chat, or video calls, adhering to cybersecurity policies becomes crucial. [28].

Instead, working from home should not compromise an organization's security standards against impersonation, malicious content, and other social engineering tactics. Even from home, where the availability of IT support might be limited, these policies guide employees, helping prevent common mistakes and maintaining

the organization's overall security posture[29]. To combat against social engineering, as well as other cyber attacks, policies are outlined. Such policy should not only cover precautions employees expected to take, but also guidence in case of a security breach[31].

2.8.4 Device Management

In today's work landscape, where many forms of communication like email, chat, and video calls are commonly used, it is crucial for employees to adhere to cybersecurity policies[28]. This is true regardless of whether employees are working remotely or in an office setting. Remote work does bring its own set of challenges, such as the limited availability of IT support[29]. Regardless of location, these policies should guide employees to avoid common mistakes and maintain security.

2.9 Technical countermeasures against social engineering attacks

Technical defences can be categorised into software and hardware solutions to boost an organisation's cyber security posture. While many of these tools are generally designed for cyber security, a subset of those measures addresses social engineering threats. Firewalls, which can be both software and hardware-based, manage incoming and outgoing traffic, blocking unwanted or suspicious activities. Antivirus software, on the other hand, actively scans systems for malicious content. These tools, when implemented correctly, enhance the general cyber security landscape but also provide a line of defence against social engineering attacks. Some of these are used actively, while some are used more passively. It can be argued that there is a distinct notion of the responsibility given to the individual workers at the company, presumably, having security in mind, so is productivity and usability.

2.9.1 Password Management

Choosing a password is easy, but remembering many of them can be tricky. In 2020, a survey found that individuals have an average of one hundred passwords per person to account for. Remembering such amounts of passwords can be tedious, especially if many of them are required to be reset regularly[32].

2.9.2 Multi-factor Authentication

Multi-factor authentication introduces an additional security layer during the authentication process, often necessitating an action on a secondary device. While this means that stolen login credentials alone are insufficient for a successful breach, it does not make such credentials useless to social engineers. An example of such an attack is the MFA Fatigue attack. In this tactic, attackers use scripts to flood individuals with a series of authentication prompts, leveraging on fatigue that eventually leads to an unintended approval. Such attacks compromise the overall security posture by capitalising on an individual's fatigueness. One of the mitigations against MFA Fatigue attacks is the integration of One-Time Passwords (OTP) for authentication approvals[33].

Chapter 3

Literature Review

This chapter provides an overview of the current knowledge on social engineering and countermeasure by doing a systematic literature review. We conduct a search for relevant literature, mostly using Google Scholar and Aalborg University Library, including cross-double-checking the literature, which incorporates databases from various publishers such as Elsevier, ACM, IEEE, Springer, and others. The search term used was "social engineering", "social engineering attacks", "social engineering in cyber security", "social engineering countermeasure", "- mitigation", "-prevention" including variations of these terms. After finding potential papers, I then recursively examined the cited literature from the references to expand the search and ensure a thorough review of the relevant literature.

As part of the inclusion methodology, with a few exceptions, the evaluation metrics considered are the following:

- Studies conducted in the last 10 years
- Venue, Journals and Conference quality, incl. H-index and Impact Score
- Number of references
- Number of citings (the number of times other paper referenced the paper)
- And lastly, an overall assessment of the quality given the paper's goal and contribution.

Apart from a single technical report, x social engineering surveys, the majority of papers considered for the literature review are scientific papers focused on prevention and mitigation against social engineering attacks, i.e. countermeasures. It follows that papers that merely mention social engineering countermeasures as

potential solutions without providing any scientific evidence or evaluation to support their effectiveness will not be considered in the literature review. As to this exception, several papers regarding social engineering and countermeasures that may be used throughout the report would be excluded from the literature view.

3.1 Cyber Security

Cyber security is a widely used term and can cover many different areas. A broadly accepted definition is the goal of preserving the three pillars of security; availability, integrity and confidentiality (CIA-Triad) of an organization and user's assets.[34][35] cyber security can also be described in different ways, depending on the viewpoint being considered, whether it is cyberstalking or -bullying on an individual or a DDos attack on an organization using means of computer networks.[34]

3.2 Social Engineering

Although there is a general consensus on the definition of social engineering, a certain level of uncertainty exists within the field, especially in the area of cyber security. To better understand the current meaning of social engineering, two papers have been selected for review: Wang et al.[36] and Hatfield[37]. While [37] investigates the theoretical aspects of the term, tracing its roots from politics through early-stage online hacking forums and into cyber security, [36] adopts a more pragmatic approach, focusing on the practical aspects of social engineering. In their paper, Wang et al.[36] address the shortcomings of the distorted perception of Social Engineering in Cyber Security (SEiCS) and argue that the lack of a coherent understanding of social engineering negatively impacts the ability to defend against it. In addition to examining the evolution of the term 'social engineering' (see Figure 3.1), the authors propose a refined definition of SEiCS that "eliminates inconsistencies and vagueness, encompasses mainstream intentions, and clarifies the conceptual boundary.

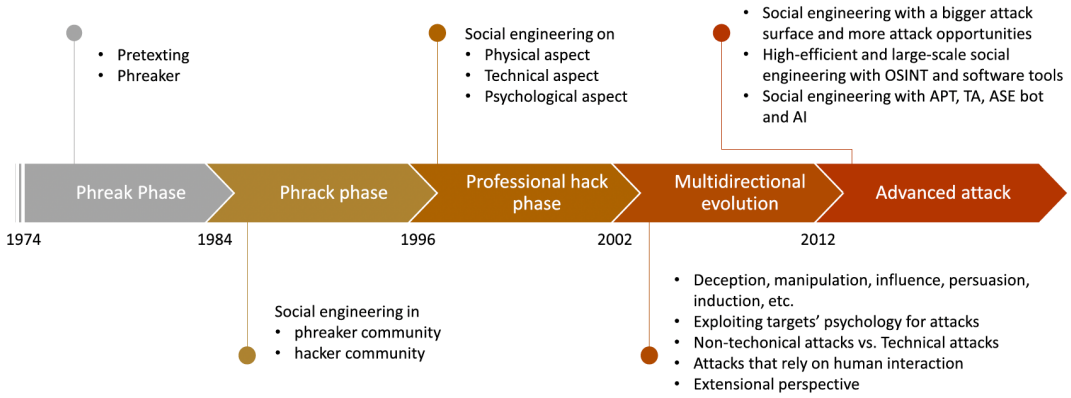

Figure 3.1: The conceptual evolution of social engineering in cyber security.[36]

The paper demonstrates its findings through five analysis tables. For example, in Table 5 of the paper, the authors present some cyber attack scenarios which do not necessarily fall under a social engineering attack but highlight the depending acts of social engineering in the same cyber attack. See Figure 3.2.

Attack	Description of non-social-engineering part	Description of social engineering part
XSS & phishing	The attacker finds website A containing a non-persistent (reflected) XSS vulnerability that enables attacker to inject client-side scripts into web pages, and then crafts an innocent-looking URL that call attack scripts (e.g. exp.js) prepared in sites B to exploit the vulnerability.	The attacker sends the malicious link to targets (e.g. some authenticated user of website A) via email, instant message, etc. to trick targets into clicking. Once clicked, the attack script exp.js in site B is triggered to load in victims' browser as if it originated from website A and to execute malicious functions such as cookie theft and privilege escalation.
CSRF & spear phishing	When a web server A is designed to receive a request from a client without sufficient mechanism for verifying, then it might be possible for an attacker to conduct a CSRF attack. E.g., targets are currently authenticated on website A, and credentials associated with website A, such as the user's session cookie, IP address, Windows domain credentials are saved in browser.	The attacker embeds forged request into website B in advance, and lures targets to visit website B via e.g. instant message. If the target is not logged into website A, a (spear) phishing email can be used to trick targets into logging in website A and clicking a link points to website B. Once the targets opened the website B, the forged request in website B inherits the identity and privileges of the targets on website A to perform malicious functions on website A automatically through targets' browser.
Drive-by Download & phishing	Attackers first create malicious content (e.g. a computer virus, spyware or malware) and host it on their own server or a compromised legitimate website, and then prepare the "drive-by webpage" by compromising other websites and injecting small pieces of code with download functions inside.	Attackers send a link in an email, text message, or social media post that tells targets to look at something interesting or awards on "drive-by webpage" (exploit human vulnerabilities such as curiosity and greed). Once the link opened, without requiring further victims' interaction to click or accept any software, the malware can be download and even executed in the background without victims' knowledge (exploit the vulnerabilities in the browser, plugins or system).
APT	Advanced Persistent Threat is a cyber security threat actor, typically a nation state or state-sponsored group, which gains unauthorized access to a computer network and remains undetected for an extended period.	Social engineering usually serves as an important attack stage of APT, to gain access to a physical location to enable further network attacks in initial stage, to bridge other attack vectors in the intermediate stage, etc. Stuxnet and Flame are cases in point.

Figure 3.2: Cases analysis of social-engineering-based attacks.[36]

3.2.1 Phases of Social Engineering Attacks - A lifecycle

Social engineering attacks are a type of attack that often involve a series of four phases, each serving a distinct purpose. These phases are Research, Hook, Exploit, and Exit Strategy.[38, 39, 40]

- **Research** During the Research phase, attackers gather information about the organization's structure, processes, and potential vulnerabilities. This information is then formed into a strategy best suited for the target.
- **Hook** The Hook phase involves creating a scenario or pretext that will allow the attacker to gain the trust of employees and establish a relationship.
- **Exploit** In the Exploit phase, the attacker uses the information and relationship established in the previous phases to manipulate employees into performing a desired action, such as divulging sensitive information or granting access to restricted areas.
- **Exit Strategy** Exit Strategy phase involves covering the attacker's tracks and ensuring that they are not detected or identified. Understanding these phases is critical for organizations to recognize and defend against social engineering attacks.

3.2.2 Social Engineering Attacks Taxonomy

The types of social engineering attacks range from phishing emails to phone scams, and they often involve tricking individuals into exposing sensitive information or performing actions that they shouldn't do. A way to classify social engineering attacks is, firstly, to base the types on the level of involvement of human interaction, i.e. human-based vs computer-based social engineering attacks[38].

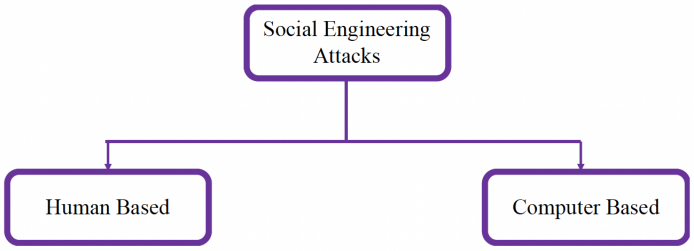

Figure 3.3: Social engineering attacks classification[38]

In this context, human-based attacks involve direct interaction between the attacker and the victim, while computer-based attacks may rely in part upon automated tools and software to carry out the attack. Furthermore, while human-based attacks are a limited number of victims due to the need for in-person interaction, computer-based attacks are limited in the technology used, such as mobile phones and computers. Computer-based attacks make computer-based social engineering

attacks easier to perform on a larger scale.[38]

Second, according to the papers by Krombholz et al.[39] and Arabia et al.[41], as an extension to the first categorization, the two types of social engineering can be perceived as multifaceted, which can be explained in the different stages of an actual attack, namely, physical, social, technical and socio-technical. These terms can, in turn, be used as terms referring to the focus of countermeasure techniques[42].

- **Physical** refers to the act of physical efforts with the goal of gathering information about a potential victim. An example of this is dumpster diving.[39, 41]
- **Social** refers to "social" parts of social engineering. These methods rely on the psychological aspects of an attack.[39] EXAMPLE??
- **Technical** refers to the attacks which are executed with electronic devices, mainly over the internet.[39, 41]
- **Socio-technical** refers to a combination of several methods used, such as bating, where a social engineer would place a USB containing malware at a point for a target to find.[39] The goal of this would be to have the victim open the USB containing malware on a work computer exposing the organisation. The baiting tactic relies both on human interference and computers.

Furthermore, given this taxonomy, the authors of [39] also mapped the relationship between the classification in line with a selection of social engineering attacks, or Channels, respectively. See Figure 3.4.

		Phishing	Shoulder Surfing	Dumpster Diving	Reverse Social Engineering	Waterholing	Advanced Persistent Threat	Baiting
Channel	E-mail	✓			✓		✓	
	Instant Messenger	✓			✓			
	Telephone, VoIP	✓			✓			
	Social Network	✓			✓			
	Cloud	✓						
	Website	✓				✓	✓	
	Physical	✓	✓	✓	✓			✓
Operator	Human	✓	✓	✓	✓			✓
	Software	✓		✓	✓	✓	✓	
Type	Physical		✓	✓				✓
	Technical					✓	✓	
	Social				✓			
	Socio-technical	✓			✓	✓	✓	✓

Figure 3.4: Classification of social engineering attacks[39]

Here the authors roughly recognize the two types of social engineering attacks, i.e. human-based and computer-based attacks, and label them under the category "Operator". Note that the included papers classifying the types convey the same concept but use different terms to describe it.

Thirdly, Xiangyu et al.[43], who mostly focuses on insider threats, categorize social engineering attacks according to timeframe, or attack strategy, respectively. Here, the authors explain that to make the categorization simpler to understand; they further appropriate the word usage into short-term attacks and long-term attacks. Short-term attacks consist of single event non-repeating attacks, while a long-term attack refers to a strategy that is designed to be carried out over an extended period of time where a sequence of steps is taken. In line with this definition, a technical report from Intel Security report[40] appears to align closely with the classification described in [43] as the classifications; hunting and farming. In short, the technical report states that hunting (short-term attack) uses minimal interaction with the target and typically involves a single encounter ending the relationship after the goal is acquired. In comparison, farming (long-term attack) seeks to establish a long-term relationship with the target, with the intention of

gradually extracting valuable information over an extended period of time.

3.2.3 Social Engineering in the COVID-19 Era

The COVID-19 pandemic has notably influenced the cyber security landscape, with a specific emphasis on social engineering attacks. During this period, various work arrangements, such as working on-site, remotely, or following bring-your-own-device policies, may unintentionally undermine the security protocols within an organization.[44, 45]

For this purpose, four studies as selected for review. A multivocal literature review by Hijji et al.[44], a paper by Lallie et al.[45], and a study conducted by Hoheisel et al.[46].

Firstly, in [45], the authors investigate the impact of the COVID-19 pandemic on cyber-attacks, including social engineering attacks. The study reveals that cybercriminals took advantage of global events and governmental announcements to create targeted phishing campaigns, increasing their chances of success. Table 1 in the paper enumerates 39 COVID-19-related cyber-attacks, categorized by one or more attack types. These attack-type categories include phishing (or smishing), pharming, extortion, malware, financial fraud, and hacking. According to the paper's classification and attack descriptions, the cyber attacks that are considered as social engineering attacks involve phishing, pharming, and extortion. Of the 39 attacks, more than 80% partially or fully utilized social engineering techniques. Ultimately, it is important to note that the paper provides references for each of the listed cyber-attacks. In order to accurately evaluate the level of social engineering involved in these attacks, a thorough examination of each reference would be necessary.

Second, [44] is a multivocal literature review (MLR), a type of systematic literature review(SLR), that combines scientific papers with grey literature (GL). In the context of social engineering, grey literature refers to testimonies from practitioners and professionals within the field as well as blogs and web pages. Doing an MLR, as appose to SLR, a study by Garousi et al.[47] was conducted in order to examine the benefits of including sources of material not presented in scientific studies within the field of software engineering. The study found that, in some cases, excluding the GL would neglect "a major pile of experience and knowledge..". The authors of [44] suggest that, in order to study social engineering techniques used during the COVID-19 pandemic, the contribution of GL is valuable. The study revealed that, given the paper's six research questions and their literature research methodology, practitioners are significantly more active in offer-

ing social engineering-based cyber security solutions in the time of the pandemic. See Figure 3.5.

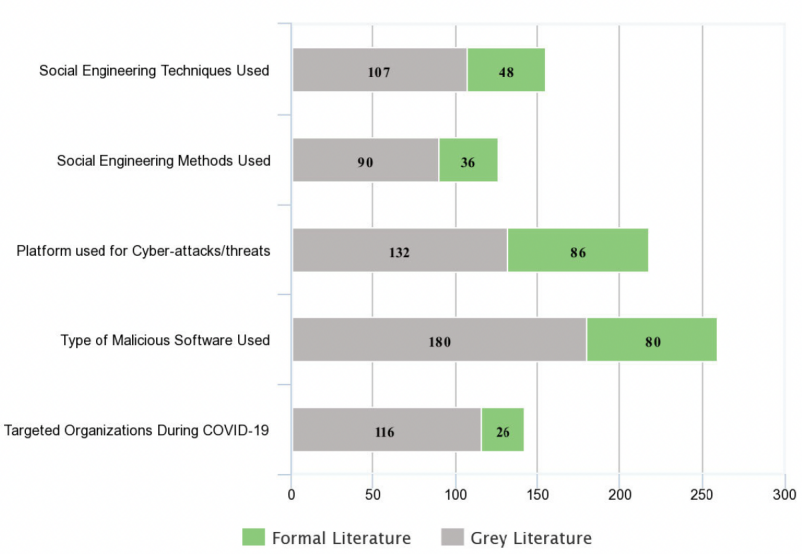

Figure 3.5: Frequency of research questions in formal and grey literature sources[44]

Furthermore, the paper identified several challenges and potential attack vectors given social distancing, working from home, and other changes in behaviour during the pandemic. These new conditions have required adaptation from organizations in order to maintain their security posture. From this, the authors pose the challenges into seven categories and attempt to provide recommendations. However, these recommendations are seemingly basic or straight foward recommendations, or so-called tips, respectively. As well as mentioning non-specific countermeasures recommendations, such as training and awareness, AI, and using Big Data for anomaly detection.

In summary, it is important to cautiously consider the potential long-lasting impacts of the working trends observed during the COVID-19 pandemic. Doing Google Trends queries of the term "work remotely"[1] and "working remotely"[2] in the past five years worldwide indicates a significant spike in 2020, followed by a reduction but still maintaining a higher trend than before COVID-19. This suggests that, while it is not certain, there might be a lasting shift in work policies towards remote work, which could potentially maintain the vulnerabilities associated with

[1]https://trends.google.com/trends/explore?date=today%205-y&q=work%20remotely&hl=en
[2]https://trends.google.com/trends/explore?date=today%205-y&q=working%20remotely&hl=en

social engineering attacks. However, further research and observation are needed to understand the long-term implications of these trends on cyber security fully.

3.2.4 Human Vulnerabilities

A part of social engineering attacks lies in the acts of psychology and persuasion techniques. These manipulation techniques can be significantly associated with social engineering described by social engineering practitioners Kevin D.Mitnick[2] and Christopher Hadnagy[28] and psychologist Robert Cialdini[48]. In fact, the foundation of [48] which is presented as six principles[3] are referenced in the social engineering literature, including the study by Wang et al. and [28, 2]. The first two mentioned books, however, focus on the practical aspect of utilizing social engineering, giving detailed examples of social engineering which is generated from experience.

To shed light on which and why social engineering succeeds in praying on human nature, Wang et al.[49] attempt to investigate the psychological aspects of social engineering attacks. They highlight the importance of understanding human vulnerabilities, effect mechanisms, and attack methods as the three core entities; effect mechanism, human vulnerabilities and attack methods. Furthermore, they examine over 30 effect mechanisms across six aspects, such as persuasion, social influence, emotion and decision-making, and trust and deception, to provide a view of the psychological factors contributing to these attacks' success. The study also discusses over 40 human vulnerabilities in areas such as cognition and knowledge, behaviour and habit, emotions and feelings, and human nature, as well as personality traits and individual characters. These entities are described as a combined view of a schematic diagram in paper's Figure 2. One of the paper's takeaways is the 16 case study where they map the correlation between core entities. Although there are 16 case scenarios, a few as been selected, see Figure 3.6.

[3]Reciprocity, Commitment & Consistency, Social Proof, Liking, Authority, and Scarcity

No.	Social Engineering Scenarios Description	Effect Mechanisms	Human Vulnerabilities
1	**Pretexting**. The attacker requests classified information by pretending to be a cable splicer and pretexting that he is wiring two hundred pair terminals for police. Who would want to refuse a little help to a company man coping with that heavy-duty assignment? She feels sorry for him, she's had bad days on the job herself, and she'll bend the rules a little to help out a fellow employee with a problem.	Social responsibility norm and moral duty, similarity & liking & helping, emotions and feelings influence decision-making, ELM, IDT, factors affecting trust.	Sadness, sympathy, the desire to be helpful, agreeableness, kindness and charity.
2	**Vishing and Pretexting**. The attacker pretends to be a new employee and convince the targets that he will suffer greatly if the request is not granted. E.g. request the technical support (e.g. Paul) to reset the password of certain account to deal with an urgent task, and further ask a VPN to access from outside.	Foot-in-the-door, impression management theory, ELM, two routes to persuasion, IDT, cognitive dissonance, emotions and feelings influence decision-making.	Guilt, the desire to be helpful, friendliness, credulity.
3	**Vishing and Pretexting**. The attacker calls a staff of the technical support department to say that the CEO authorized his requesting an urgent VPN channel for a project presentation in another city, and further tells he / she that other staffs did this before, such as Paul.	Source credibility and obey to authority, diffusion of responsibility, bystander effect, deindividuation in group.	Fear and dread, neuroticism, the desire to be helpful, friendliness, credulity.

Figure 3.6: Cases study of 16 social engineering attack scenarios to illustrate the application of the conceptual model.[49]

The paper does not result in specific quantitative findings, as the primary focus is the conceptual exploration of social engineering attacks. Instead, they provide a model to describe how social engineering attacks work and take effect, highlighting the complex interaction of the psychological factors that are central to these cyber security threats.

3.2.5 Open Source Intelligence

Open source intelligence, or OSINT, in social engineering, refers to data which can be used maliciously as a bridge into social engineering attacks. In a paper by Huber et al.[50] an experiment was conducted using an automated chatbot on social media, specifically Facebook, to collect information from individuals, in this case, students. The chatbot, called ASE bot, operates by mapping potential victims from a specific organization's network who match predefined criteria. The ASE chatbot specifically targets profiles with "open" privacy settings and terminates the process if the number of matching profiles is insufficient. This approach holds significant value, as it eliminates the need to establish a relationship with the target manually. The results were somewhat convincing, as not all participants could determine that the chatbot was not human. Although this paper was published in 2009, the concern remains relevant. As AI chatbots, including models like ChatGPT, continue to improve and become more sophisticated, there is a potential for misuse for malicious social engineering efforts.
From this, as an inspiration, and in an effort to expand this concept, Edwards et al.[51] did a similar experiment. Their methodology was based on guidance and in-depth interviews with social engineering experts. Here the authors presented an approach for identifying employees of an organization among numerous connected profiles in online social networks, using only publicly available data. This method involves gathering employee information based on the organization's social media footprint across multiple online services including LinkedIn, Twitter and Face-

book. Instead of using an invasive chat procedure, the system passively expands its target profiles by conducting broader searches for the individual's online presence. The authors suggest that, as a form of penetration testing, an automated social engineering scan across various platforms to detect footprint and OSINT is recommended.

3.3 Social Engineering Countermeasures

As discussed in Chapter 2, social engineering attacks pose a significant threat to organizations and individuals by circumventing security policies, exploiting vulnerabilities, and employing psychological manipulation and persuasion techniques to deceive and exploit human vulnerabilities. Countermeasures have been proposed and developed to mitigate these attacks preventing access to sensitive information. The papers reviewed here go beyond classifications made in Section 2.7.

3.3.1 Social Engineering Awearness

Several types of security awareness training are the most popular response for social engineering mitigation.[51][52]. Newer awareness techniques involve developing simulation e-platforms to enhance employee learning and awareness. A type of platform is a gamification platform where the employee is exposed to social engineering principles. Because social engineering poses an unconventional type of cyber attack to organizations, it is essential to ensure that every individual, from top management to support staff, including the cleaning personnel, maintains good cyber hygiene. As organizations become increasingly dependent on information systems, social engineers employ more technologically advanced methods to pose a significant threat[52]. Training and awareness cover both the understanding of social engineering and overall cyber hygiene.

Instead of depending only on human actions or technical methods, Heartfield et al.[53] proposed the "human-as-a-security-sensor (HaaSS)" concept to take advantage of these skills. Since humans may be better at detecting semantic deception than automated security systems, the proposed system incorporates human sensing with machine learning. The experiment participants were exposed to several types of emulated attacks, including fraudulent websites, phishing emails, malicious social media links and malicious Google Drive files. The results of the experiment were based on the effectiveness of HaaSS compared to current technical defence systems that claim to offer technology, such as Google services, Windows Defender, and antivirus software, to safeguard users from phishing and social engineering attacks. A noteworthy aspect of the experiments is that, in some cases, the

participants were exposed to two social engineering attacks depending on the first exposure. The experiment showed improved results on using HaaSS compared to other technical measures, especially browser features and antivirus features. Where the participants practically scored better on every count using HaaSS, some attacks proved to be the most challenging to detect, specifically spear phishing email attacks. The experiment's limitations are that the number of participants is low at 31 and that the participants were purposely aware of the threats in a controlled environment lacking the realistic scenario factor.

3.3.2 Technical Social Engineering Defense

Machine learning has proved to be a good tool for filtering and detecting fraud patterns. However, machine learning can indeed be helpful, but only to a certain level. For different kinds of social engineering attacks where phishing is part of, e.g. fraudulent URLs, link or file naming, email spam or general deceiving, etc., a 100% detection rate is nearly impossible due to its human-centric nature. In a paper by Sahingoz et al.,the authors explored the challenge of detecting phishing attacks in real-time by proposing an anti-phishing system that uses seven different classification algorithms and natural language processing (NLP) based features. The proposed approach achieves a 97.98% accuracy rate for detecting phishing URLs using the Random Forest algorithm with only NLP-based features. The authors recommend developing a new subsystem for shorter URLs that require additional data, such as the number of visitors and domain registration dates, to detect types of phishing attempts.[54]

3.4 Strengths and Limitations of Existing Literature

To the best of our knowledge, considering the topics that universally fall under the term social engineering, many papers exist on the topic to the point of being well-researched. However, academic research in the last decade seems to be limited in scope and type. While conducting the literature review with carefully considered evaluation criteria, some papers from less renowned journals or conferences, with fewer references, were still cited positively by other research works. Though these papers did not qualify for the current review, their potential insights and perspectives should not be overlooked. Additionally, many publications, surveys, and literature reviews rely on sources dating back to before 2010, which may still be relevant but, nevertheless, should be noted.

The majority of the research found is quantitative work as opposed to qualitative work. Given the practical nature of social engineering attacks, explored in

Chapters 2 and 3, more qualitative research should be considered.

The correlation between individual knowledge and skills within technology and cyber security threats, especially within social engineering, and their application in routine scenarios is not sufficiently explored in existing research. A deeper understanding of how people's personal backgrounds, experiences, and cognitive biases influence their susceptibility to social engineering attacks could provide valuable insights for developing more effective prevention and mitigation strategies.

Social engineering is a multifaceted challenge that cannot be solved by relying solely on technical countermeasures, such as email filtering and antivirus software. While some conceptual research exists on these technical strategies, the actual solutions themselves are often considered proprietary and not publicly available due to their commercial value. Consequently, a more in-depth examination and improvement of the theoretical basis for these and other measures are necessary to address the complexities of social engineering threats. Some reviewed papers only provide introductory overviews of potential countermeasures without offering actionable solutions. While these papers are valuable starting points for understanding the wide range of potential strategies, specific and effective solutions are needed to combat the challenges posed by social engineering threats.

Lastly, phishing attacks are often the focus, which can be perceived as organisations' main concern. Several papers address information security policies in general without directly discussing social engineering or even mentioning it. In some cases, these papers mention phishing attacks at best, which highlights the need for more research, specifically focusing on social engineering threats and countermeasures. However, it should be noted that phishing relies on some form of deception, and there are many types other than email phishing. Phishing seems to be the go-to term for social engineering attacks, but from the literature, it becomes clear that phishing encompasses various other typical social engineering concepts.

The challenge in social engineering is that no one-size-fits-all solution exists, therefore, additional qualitative approaches, such as interviews and case studies and observations, may be appropriate in order to learn the human factors at play sufficiently and devise tailored solutions that can be integrated into everyday practices and training programs.

In conclusion, the literature on social engineering indicates a complex landscape, heavily oriented towards quantitative research and email-phishing attacks. Despite extensive studies, there remains a significant gap in the understanding of human factors and individual decision-making. The limited scope of qualitative insights and the proprietary nature of technical solutions have further restricted a comprehensive grasp of social engineering's multi-faceted challenges. This review

emphasises the urgent need for a more nuanced approach that integrates qualitative methods with quantitative insights.

By examining the human dynamics in social engineering and broadening the focus beyond phishing, researchers and professionals can work towards creating more targeted strategies. These strategies could provide enhanced protection for both individuals and organisations.

Chapter 4

Methodology

4.1 Research design and approach

As discussed in Chapter 3, based on the literature review's observation, most studies on social engineering primarily use quantitative methods. However, to deeply understand individuals' knowledge about social engineering and existing countermeasures, as discussed in Chapter 2, a qualitative approach is more fitting. Such an approach delves into the nuances of backgrounds, experiences, and knowledge that may influence susceptibility to social engineering attacks. Furthermore, it allows a comprehensive exploration of both individual views and organisational countermeasures.

4.2 Data collection process

Qualitative research is rooted in its potential to dig into the complexities of human experiences and insights[55]. In order to collect insightful data on social engineering and its associated countermeasures, I began exploring locations where I could obtain qualitative information. This search led towards cyber security professionals, experts, employers, and other relevant individuals. When examining social engineering countermeasures, it is essential to consider the interplay between individual behaviours and organisational strategies.

Initially, the project's vision of collecting data entailed structured interviews and conducting experiments. However, as the research progressed, the challenge was not the limitation of the data collection method but the need for respondents willing to participate. This unforeseen obstacle required a change in approach, and as a result, the data collection for the research adapted, evolving its methods

to suit the circumstances.

4.2.1 First iteration - First Contact

For selection and contacting companies and organisations, those that are considered well-established were a priority for this study. While most interested companies are welcome, larger organisations in industries of financial services, technology, healthcare and other sectors with a higher risk of social engineering attacks will be given preference. This approach was intended to provide insights into the challenges and countermeasures organisations face that are likely to be targets for social engineering attacks.

Organisations and companies were contacted through various channels such as telephone, LinkedIn, and email to initiate a conversation and assess their willingness to be involved in the study. When a contact within the organisation was initiated, preferably a high-positioned individual, a brief introduction to the research project, its objectives, and expected outcomes was provided to help potential participants understand the purpose and scope of the study. To further convince potential participants of any kind of collaboration, the conversation started with high-stakes experiments, including experiments with various social engineering attacks and follow-up interviews. However, during the interaction, as the interest decreased, the pursuit of collaboration was minimised if not declined. This was the case with an IT security manager at a large bank, which will remain anonymous. Initially, the prospect of an interview was on the table, which developed into written interview questions. This led me to the next phase of pursuing other non-subjective stakeholders.

4.2.2 Second Iteration - Objective Insights

After encountering challenges with directly engaging well-established companies, the research shifted its focus towards stakeholders who could provide objective insights without the constraints of company partnerships. Instead of focusing solely on those within corporate security, the search expanded to experts knowledgeable in social engineering and its countermeasures. This category ranged from vendors offering solutions addressing social engineering threats to recognised experts working with or studying these techniques.

Various channels were leveraged to reach out to these individuals. Online platforms, especially LinkedIn, became a primary method for connecting with experts. When it came to vendors, contact was primarily initiated through the contact information provided on their official websites. The objective here was to gather insights unfiltered by organisational policies or security concerns.

Among the individuals contacted, Christopher Hadnagy stood out. As the founder and CEO of Social-Engineer LLC and a known authority in the social engineering sphere, his perspective was of great interest. Despite the reluctance from previous encounters with other organisations, Christopher Hadnagy generously provided an extensive interview. This encounter brought a fresh perspective and contributed substantially to this study's data depth.

However, despite the success with Chris, it became evident that relying solely on a single expert interview might not provide the breadth required for a complete understanding. While Christopher Hadnagy's insights were profound, such experts' pool was limited. This realisation led to the following research phase, focusing on broader data collection through conducting a type of survey. As part of this project, the transcript of the interview can be found in appendix B of the project.

4.2.3 Third Iteration - Survey

In pursuit of gaining deeper insights into social engineering practices within companies, an alternative approach was adopted: engaging directly with individuals employed within these organisations. The objective was to tap into their firsthand experiences and perceptions related to cybersecurity measures.

For the distribution of the questionnaire, to maximise the reach and impact of the questionnaire, it was strategically shared on LinkedIn's professional networking platform. Notably, Mr. Christopher Hadnagy, a well-known figure in the social engineering spectrum, generously shared the questionnaire on his LinkedIn profile, significantly expanding its visibility across relevant professional individuals. Furthermore, to ensure some respondents, the questionnaire was also distributed directly to selected individuals who met specific criteria for participation. These criteria entailed working in an environment where social engineering attacks can happen, meaning that individuals working as chauffeurs or at cash registers were deemed not appropriate as respondents due to their roles being less likely to be targets of social engineering attacks.

In order to capture the field-level experience and perception of social engineering, the questionnaire was designed accordingly, focusing on several areas of the individual's:

- background
- self-assessment of their cybersecurity awareness
- individuals' encounters with social engineering attacks

- understanding of both non-technical and technical countermeasures
- feedback mechanism
- view on AI in social engineering

The questions ranged from assessing their familiarity with social engineering tactics to understanding the training and tools provided by their respective organisations to counter such threats. The primary intent of this iteration was to collect individual insights, providing a view of the present landscape of social engineering challenges and the measures adopted to combat them. In the next chapter, I will provide a detailed analysis of the questionnaire design, rationale for question choices, and potential limitations that may affect data interpretation. This will offer a broad view of the data collection process and its validity. As part of this project, the survey responses can be found in the link[1].

4.2.4 Fourth Iteration - Last Interview

This study's fourth and final data collection iteration introduced an unexpected yet valuable component to the research methodology. As an extension to the second interview in the second iteration, this involved an interview with Chris Kayser, the Founder, President, and CEO of Cybercrime Analytics. The connection with Chris Kayser was established during the questionnaire distribution process. He showed interest in the research project, followed by a request to participate. This development led to the scheduling of the interview, creating an opportunity to enrich the research with his insights. Without extensive exploration of the conducted survey, a small observation from the survey was presented to C. Kayser in an attempt to discuss the results. As part of this project, the audio can be found attached to the project.

Summery

The research project aimed to adopt a qualitative approach to explore social engineering and its countermeasures, varying from the more common quantitative studies. Data collection underwent several iterative adjustments due to challenges in securing participants.

Research Design: A qualitative method was chosen to examine both individual behaviours and organisational strategies in social engineering.

[1] `https://docs.google.com/forms/d/17C5AAnG__x2Ddi-A4QhyyiSOqWZp3Clbb7aLB5pMgkE/edit#responses`

Data Collection: Initial efforts focused on well-established, high-risk sectors but shifted due to low participation rates.

The research methodology unfolded in a series of planned iterations, each designed to explore a different facet of social engineering and its countermeasures. These iterative steps ensured a multi-dimensional inquiry that adapted to challenges and opportunities encountered during the study.

- The first iteration targeted large companies, especially in high-risk sectors.
- The second iteration shifted focus to social engineering experts for objective insights.
- A third iteration involved a survey aimed at professionals to gather individual insights.
- A fourth iteration included a second expert interview with Chris Kayser.

Despite challenges, the research gathered helpful insights into social engineering countermeasures from diverse sources. In the next chapter, the interview and survey are presented and discussed.

Chapter 5

Interview and Questionnaire

The following chapter provides insight into the rationale behind the questionnaire questions, aiming to offer a perspective on the data collection process.

5.1 Bank Written Interview

Assessment and Countermeasure Strategies

Questions 1-3 ask about the bank's methodologies and challenges in identifying and mitigating social engineering vulnerabilities in order to get a view of the bank's strategic and operational approach towards managing threats. This entails understanding the technological tools employed to counteract social engineering threats and how these technologies are integrated within the bank's overarching cybersecurity strategy. Additionally, it seeks to learn the decision-making process behind choosing methods and tools to combat social engineering, exploring the influences and criteria that guide these choices.

Training and Employee Development

Questions 4 and 5 aim towards training and employee development, delving into how the bank considers an employee's previous IT experience in shaping their training in social engineering countermeasures and exploring the challenges and advantages of training individuals from diverse backgrounds. The bank is also questioned about the structure and customisation of social engineering awareness training for different employees, exploring whether the training is adapted based on roles, IT experience, or background.

Device Management and Remote Work

Question 6 addresses device management and remote work, exploring the bank's strategies and policies for device management, especially in situations where employees have more flexibility in how they use their devices, such as remote working and BYOD and policies, and how these strategies address social engineering threats like compromised devices or stolen credentials.

Effectiveness and Feedback Mechanisms

Question 7 aims to understand effectiveness and feedback mechanisms, seeking to comprehend how the bank measures the effectiveness of its countermeasures against social engineering and whether there is a feedback loop to refine and improve existing methodologies continually.

Adaptation to Evolving Threats

Question 8 explores the bank's adaptability and foresight regarding the evolving nature of social engineering tactics, such as deepfakes and vishing, and how it plans to modify its countermeasures in the coming years while also being watchful of emerging trends and threats in the field of social engineering.

5.2 Interviews

5.2.1 Chris Hadnagy Interview

Intro & Definition of Social Engineering

At the beginning of the interview, the scene is set. Setting mutual expectations ensures that both parties are aligned on the objectives, minimizing misunderstandings. It also sets the tone for a focused and productive interview. This is followed by an introduction to the interviewee to explain his background and his founded company, including the methodologies used in his company.

At the beginning of the interview, the scene is set. Setting mutual expectations ensures that both parties are aligned on the objectives, minimizing misunderstandings. It also sets the tone for a focused and productive interview. This is followed by an introduction to the interviewee to explain his background and his founded company, including the methodologies done of his company.

The first set of questions regarding social engineering was the development of social engineering attacks. As previously discussed in Section 2.2, regarding the elusiveness character and deployment tactics of social engineering attacks, a question is asked to shed light on the categorisation of past and present social engineering attacks. Knowing the interviewee's definition of social engineering can offer a foundational groundwork for the entire discussion. It can provide insights into whether the interviewee views social engineering more as a technical issue, a human issue, or a blend of both, which in turn can influence the conversation. Additionally, the question explores "legacy" social engineering attacks assessing whether they remain a relevant threat or have become obsolete, while also inviting current and newer forms of social engineering threats.

Insights into Social Engineering Countermeasures

The next questions explores organisations' ability to effectively customise its countermeasure operations based on experience. These questions take basis from the review of social engineering countermeasures outlined in section 2.7 as well as knowledge gathered from the written interview with the bank.

Strategies for Social Engineering Countermeasures

Questions 4 through 8 collectively examine the strategies organizations deploy against social engineering threats. Specifically, Question 4 explores how top organizations decide on their countermeasures, while Question 5 delves into the tailoring of awareness programs to cater to specific employee roles and threats. Question 6 investigates the influence of diverse educational and professional backgrounds on the effectiveness of training. The emphasis of Question 7 is on the significance and best practices of a feedback loop in refining countermeasures, and Question 8 contemplates the role of employees' general cyber hygiene in shaping countermeasure strategies.

Metrics of Effectiveness

Question 9 uncovers innovative methods organizations use to move beyond traditional metrics and measure the effectiveness of their social engineering countermeasures.

Navigating Transparency in Social Engineering

Question 10 seeks ways researchers can encourage organizations to engage in more transparent dialogues regarding their security measures, particularly in the context of social engineering.

The Changing Landscape of Social Engineering

Question 11 discusses the adaptations and strategies organizations are employing to counteract social engineering threats in the context of BYOD trends and increased remote work.

Future Prospects and Threats

Question 12 explores the potential implications of technological advancements, such as deepfakes and AI, on the evolution of social engineering threats and defences. The question delves into how organizations might need to proactively adjust their countermeasures in the face of these emerging challenges.

Summery

Both interviews were initially expected to follow the same set of questions. However, due to time constraints on the part of C. Kayser, the second interview was slated to last only thirty minutes, requiring fewer questions. The first interview with C. Hadnagy proceeded in a structured manner, allowing for a systematic walkthrough of the entire set of questions. In contrast, the conversation with C. Kayser took a more conversational tone. While this approach prevented in-depth exploration of specific questions, it yielded valuable insights through the natural flow of the conversation. Unexpectedly, the dialogue with C. Kayser extended beyond an hour, resulting in an open-ended interview. Seeing the interview become dynamic made it challenging to anticipate the interview's conclusion. Nevertheless, the conversation continued, addressing follow-up questions as they emerged, sometimes touching on casual topics.

As a result, each additional question felt like it could be the last, adding a unique dynamic to the discussion. Although the interviews differed in the end, both were valuable.

5.3 Questionnaire Questions

Demographic Information

Questions 1-5 are general demographic questions about the respondent, such as gender, age, education, and industry. This helps in segmenting the data later to determine if awareness or experiences vary based on these factors.

Own Assessment of Their Awareness

Questions 6-8 asks the respondent about their own perception of their competence with IT, their awareness of cyber security and finally, their understanding of social engineering. This provides a baseline understanding of the respondent's foundational knowledge in the domain. It can be argued that answering to their understanding of social engineering corresponds to their understanding of cyber security and, accordingly, their understanding of cyber security correspondence to their IT competence. Albeit, in practical scenarios, an individual with high IT competence may not necessarily have a high cyber security awareness. Some respondents might overestimate their competence or awareness in certain areas, believing that they know more than they do. Conversely, highly knowledgeable respondents might underestimate their expertise, assuming that what they know is common knowledge.

Social Engineering Countermeasures & Experience

Questions 9-17 revolve around social engineering awareness and experience and probe respondents' exposure to and knowledge of social engineering attacks in their work environment. Questions range from first-hand experiences of incidents to the effectiveness and inclusivity of organisational countermeasures and training initiatives. In order to gain some insight in organisation countermeasures, they were asked this. If respondents answer "yes" to whether they are aware that their organisation has any countermeasures for social engineering, they are prompted to describe, in their own words, the security measures or protocols implemented at their workplace. Additionally, they are questioned about their perceived adequacy of these strategies, the existence of tailored approaches, and the availability of feedback mechanisms within their workplace. Thus, these questions aim to gain insights into both the individual and organisational aspects of social engineering threats from the employee's viewpoint.

AI and Deepfake Awareness and Concern

Questions 18-20 focus on the rising threat of AI technologies, emphasising deepfakes. Respondents are asked about their familiarity and concerns about AI and deep fake technologies, which can be utilised for impersonation. If the respondents say 'yes' to being concerned, then a follow-up question will allow them to elaborate. This can give insight into understanding how aware respondents are of AI in social engineering and gives a snapshot of how ready and proactive general workplaces might be.

Desire for Training

Question 21 asks about respondents' interest in additional training on social engineering attacks, revealing attitudes towards existing knowledge.

5.4 Data Collection Take-Aways

5.4.1 Written Interview With Bank

Despite the limited and cautious responses, gaining insights from a bank on social engineering countermeasures was valuable. The bank uses threat modelling for vulnerability assessments to balance cost, user-friendliness, and security, including for social engineering vulnerabilities. Training is tailored, with developers receiving specialized modules while others get foundational training. Specific security measures remained undisclosed, but the bank's training is based on market demands, ensuring a diverse approach. Effectiveness is observed through various metrics, some of which are confidential. With the BYOD trend, the bank retains strict device security control. The rise of AI-enhanced social engineering tactics has the bank on high alert, prompting continued exploration of countermeasures.

5.4.2 Expert Interview

Both interviews were intended to be conducted with the same set of questions. However, due to time constraints, the second interview was shortened, necessitating a reduction in the number of questions. Despite these differences, both interviews proved valuable. The first interview, with C. Hadnagy, followed a structured format where each question was posed sequentially. In contrast, the interview with C. Kayser evolved into a more conversational exchange for the most part.

5.5 Data analysis approach

In this research, the data analysis approach is straightforward and interconnected. It starts with a brief review of a written interview, setting the initial context. Next is an interview with a social engineering expert. Insights from this discussion guide the design of a following questionnaire, ensuring it's informed and focused. The questionnaire is central to the research, collecting a wide range of views and responses. Its results highlight both common themes and differing opinions in the gathered data, providing a richer understanding of the subject. The final step is an interview with a cybercrime expert. This conversation offers a closer look at

the questionnaire results, adding an extra layer of insight and ensuring the findings are sound and reliable.

Throughout the analysis, each piece of data is considered in relation to the others, creating a clear, complete, and connected picture of the research topic. This method guarantees that the final conclusions are not just detailed and thorough, but also dependable and strong.

Chapter 6

Results

In the this chapter, an examination of the questionnaire responses and the insights derived from the interviews will be presented. These findings collectively contribute to a thorough understanding of social engineering practices, countermeasures, and the dynamic cyber threat landscape.

6.1 Insights in Current Attack Types

6.1.1 What is Social Engineering - Today

Looking into social engineering requires exploring its very definition, or at the least having a common ground, which is a task that invokes diverse perspectives, even among experts in the field. How social engineering is defined and understood impacts and shapes how individuals approach and respond to it, both as a potential threat and a tool. With this in mind, the following discusses insights from two experts, shedding light on the multifaceted nature of social engineering.

In an open question regarding what social engineering represents today, as opposed to its traditional or widely recognized definition, and its misconceptions, both C. Kayser and C. Hadnagy provide varied perspectives. While C.Kayser says:

> "Social Engineering, to me, is simply getting people to do what you want them to do."

Reacting to a similar self-made expression, C. Hadnagy adds:

> "I don't really like that definition because I think that when you word it that way, you're saying that social engineering is always negative ...it's not manipulating them(his kids), it's using influence and good communication tactics to get people to do things so to me."

Through the usage of everyday analogies, such as eating healthier and parenting, C. Hadnagy explains that the misconception of social engineering is that it is exclusively negative or a method of attack, but in fact, a communication method which can be used for good.

6.1.2 Endurance of Traditional Social Engineering Attacks

In seeking insights into what type of attacks dominants, C. Hadnagy highlights the notion that the core mechanisms of attacks today have remained relatively consistent. With the exception of the more recent smishing, established methods like email phishing, vishing, and impersonation continue to prominently occupy the social engineering landscape. Here, C. Hadnagy explained:

> "The four base attacks are the same, minus smishing, and they've been the same since the beginning of man. Think about it: we didn't have email, but there were scams happening in person, and letters were sent by con men."

As previously presented in Chapter 2, the social engineering attacks were divided into platforms and execution. While C. Hadnagy acknowledges this categorization, he suggests an additional nuance. He explained that, whereas impersonation attacks are classified as a medium for an attack in one context, he interprets it as a distinct type of attack in itself rather than merely a method of execution. He further explained:

> ..when we talk about our medium of delivery, is it phone, is it text message, is it email, that's good. Impersonation could be delivered through all three of those. But it also can be a separate attack, in the sense like we see LinkedIn impersonation non-stop.

At the same time, referring back to his previous statements, the four dominating attacks are channels of communication used maliciously, indicating that the approach of conducting social engineering attacks may stay the same, in principle, but with other tools or methodologies. Thus, should another communication channel gain prominence or become widely adopted in the future, it may similarly be leveraged for malicious purposes in social engineering attacks.

When asked about the persistence of traditional social engineering tactics amidst growing digital strategies, an example is offered by C. Hadnagy, which emphasises the unexpected longevity and success of what might be considered 'old-fashioned' techniques in the digital age. On the question of legacy attacks, C.Hadnagy said:

> "I would say there are still times—and I can't believe it—that we go dumpster diving. And I find unshredded documents. This just happened recently; we were on a job. Grab a garbage bag, open it up, and there's a whole bunch of financial statements, not shredded, right?"

This statement highlights that, despite the substantial advancements in cyber security and the shift toward digitalisation, conventional methods, such as dumpster diving for sensitive information, still hold a place for social engineers. This paints a picture of a landscape where cyber security strategies must consider various attack vectors, from the digital to the physical domains. As C. Hadnagy elaborates, he indicates that such setbacks typically occur due to individuals being fatigued or tired, which, presumably, brings other similar human factors such as laziness or ignorance into play.

6.2 Choosing Social Engineering Countermeasures

In the written interview, a question was asked regarding which criteria are considered to select their methods and tools specifically for countering social engineering attacks. As a follow-up question, to add specificity, the question was further elaborated on whether they consider industry standards, benchmarks, or peer recommendations influencing these decisions. Unfortunately, it appeared that the question's phrasing was not clear enough, as the received answer did not fully address the question, stating:

> "Not quite sure I understand the nature of this question"

To get further insight into how an organisation chooses countermeasures and whether there are some accepted standard criteria, a similar question was asked C. Hadnagy. He stated:

> It's hard to benchmark this because not every company, let us say not every bank, will have the same security setup. They will not have the same protocols or educational services. So, the issue is, how do you benchmark things when there is no baseline? To set a baseline, you would have to have one bank that you say is the best on earth, and then you rate every other bank against that bank."

Here, C. Hadnagy draws attention to the significant benchmarking challenge due to organisations' diverse security and procedural landscapes. He reveals an

industry-wide difficulty in developing universally applicable benchmarks in a landscape where each organisation and its corresponding threat types vary. Hit perspective highlights a theme which is choosing countermeasures stems in assessing potential vulnerabilities.

6.3 Asessing Vulnerabilities

Questioning the bank's approach to assessing vulnerabilities related to social engineering, specifically what areas of focus and the challenges they encounter in these assessments, the bank provided the following response:

> "Such a process is multifaceted. We make use of threat modelling, in which case social engineering aspects are also assessed for the final solution. The key challenges in such assessments and decisions in terms of deciding security level are balancing Economy (cost of implementation), Usability (ease of use), and Security (restrictions, checks, tests etc.)"

The bank's usage of threat modelling highlights the importance of evaluating multiple variables - including cost, usability, and security - in developing a cyber security strategy. While offering an into a bank's approach to assessing cyber security threats, this statement is somewhat broad and opens up several avenues for exploration. Focusing on the challenge of managing cyber security as a whole within an organization, C. Hadnagy states that the problem is twofold. On one hand, he states it is common that companies need technical defence against, specifically, phishing, such as IDS (Intrusion Detection System), IPS (Intrusion Prevention System), firewall, antivirus, etc. Proceeding to explain the challenge he state:

> "You take all of that, and you need a giant budget, and this is just for phishing. We haven't even talked about vishing or smishing or anything else. So they look at this and they say, 'OK, 90% of all breaches have a phishing element to it. Let's spend most of our money there.' And sadly, this is just the nature of the world, the bad guys aren't limited by spending. If it takes them a year to infiltrate your network, they'll spend the year to do it. But our good guys, they're told, 'Here, you have $500,000. Spend it wisely.'"

So, before even beginning to open the catalogue and choose a vendor for a solution, the funding distribution needs to be shared wisely. And as he explains, he says that at the time of having allocated the money, and now choosing to vendor does not stop there. He continues his statement, saying:

> "...when you have a guy or a gal in the position who may not know everything about it, how do you know where to spend the money properly? So it's set up kind of like, I hate to say it, but it's like set up for failure, you know"

Here he expresses that how difficult it is to manage the various count to measure not only for general cyber attacks, but specifically for social engineering. Assumably, organizations which operates with counter measurements not suiting their infrastructure or employees, would wish to improve on the misalignment. From the survey, a question was asked whether there was a system in place for giving feedback on the social engineering countermeasure provided by the way place. Below is the response.

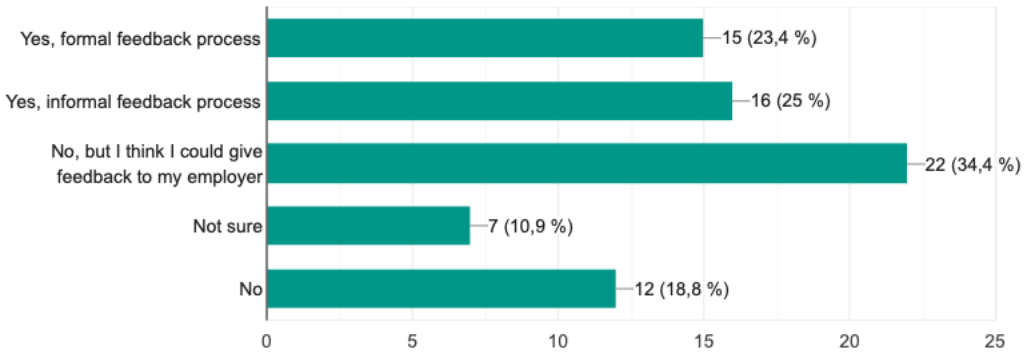

Figure 6.1: Is there a system in place for you to give feedback on the social engineering countermeasures provided by your workplace?

From the above figure, we can see that almost 23.4% of the respondents are no they can give feedback on the countermeasures, integrated into their organisations. And given that organization is ought to reconize their inability to have their workers work, should be taken into account.

6.3.1 OSINT Threshold

Open-source intelligence, which was explored in Chapter 3, has proven to be a valuable method for contacting social engineering attacks. On the matter of resilience of older social engineering attacks, as well as newer forms, instead of latching to the discussion, C. Kayser quickly expressed concern about the threshold of exposing too much information publicly. Specifically, he mentions social media as the primary example, though not exclusively, it seems. He continues describing how information can be used against individuals in various scenarios through illegal information sharing, for example, on the dark web. Here, he states:

> "There's so much data out there on everybody now. I know I've been pawned. I know that I've got personal information all over the dark web. You do? I do. Everybody does. And that information can be used indefinitely ... there's so much more information available to the cybercriminals. People are so much more engaged in technology. I think they let their guard down. I think a lot of cases, they don't care. And, of course, if they've never been hacked directly, it's the old this will never happen to me because I'm not important enough."

It has been reported that collecting data points and linking them together creates something tangible. In 2000, it was reported that, based solely on their 5-digit ZIP code, gender, and date of birth, 87% of the U.S. population, or 216 million out of 248 million people, were found to have reported characteristics that likely made them unique [56]. In contrast, He ends his explanation by saying:

> "We are the problem. We remain the problem. It's us. It's the human element, the human factor, and we have to be much more protective of our information and what we do."

6.4 Evaluating Success

It's questionable when to know if you're overall strategy works. From the bank statement, which said that they used click ratio to determine success, C. Hadnagy disagrees, stating:

> So think about this, this is my personal issue with a lot of phishing SaaS', software as a service.They focus only on click ratio, but click ratio is useless by itself.

Being in the field, C. Hadnagy expresses that reporting is a major factor in both the overall security structure and overall success of evaluating the awareness training. He mentions that the behaviour of individuals reporting ability is powerful, stating:

> reporting ratio, that tells you if your education is making a dent. So now what I say is that for phishing you grab four stats: those who clicked and didn't report, those are your worst employees. They clicked the email and didn't do anything to save the company. Those who clicked and did report, okay they did something bad, but then they realized, 'Oh, I shouldn't have clicked that,' and they reported it. That's good. Those who did not click and did not report, those people didn't

> do anything. Maybe they're not reading their email, it goes right to junk, they just deleted it. That's not good. You want to change that behavior. And those who did not click and did report, those are your star employees.

C. Hadnagy believes that when measuring the success of the current social engineering effort, one should look at the behaviour of the employees rather than the numbers.

6.5 AI Usage For Social Engineering Attacks

In an effort to probe into the topics of using AI or machine learning in social engineering attacks, a subject which C. Hadnagy expresses that he actively is researching with other researchers, a question regarding this was asked. He expresses that machine learning does nothing for the defence of social engineering attacks, answering wether we are ready of AI in the current state:

> "We're not ready. The advancements in AI have happened so fast, and there's no regulation; no one's regulating it. They're pumping stuff out. The other day, I was messing with ChatGPT-4. Because I'm a public figure, I said, "Hey, you know, tell me about Chris Hadnagy." So it told me about me. And then I said, 'Write me an email that will get Chris to come to my Russian dating site.' It wrote a really beautiful phishing email."

With today's technology with automation bots, it has become increasingly easier to conduct social engineering attacks than it ever has been. In his following statement he state:

> " ..but I said write me a marketing email that will get Chris to click on my Russian dating site. Beautiful, beautiful. Oh, that was wonderful, right? I'm thinking, "Crap, that was really good," and it did it in seconds."

C. Hadnagy raises concern about bots taking social engineering to another stage. In the same question, regarding vishing technology, he replies that, even after 20 years in the field, he would not be able to distinguish whether voice is real or machine learning generated. This indicates that, while he is not a voice expert of such kind, if an individual who has experience in vishing training and simulations believes that current countermeasures can protect from machine learning, this ought to be recognized.

Chapter 7

Conclusion

7.1 Recommendation

7.1.1 Evaluating Success Indications

Measuring the effectiveness of a cybersecurity strategy, especially against social engineering, requires more than just looking at click ratios. As C. Hadnagy highlights, many phishing Software as a Service (SaaS) solutions only focus on this metric, which he finds insufficient. Instead, he emphasises the importance of reporting. According to Hadnagy, there are four categories to consider:

- Those who clicked and didn't report (the most concerning)
- Those who clicked and reported (realising their error)
- Those who didn't click and didn't report (possibly overlooking emails)
- Those who neither clicked nor reported (the ideal behavior).

The behaviour of employees is a more accurate gauge of a strategy's success than mere numbers.

7.1.2 OSINT & Human Risk

Open-source intelligence (OSINT) is a potent tool for executing social engineering attacks. C. Kayser highlights a significant concern regarding the excessive public disclosure of personal information, particularly on social media platforms. He illustrates the gravity of the issue by pointing out that most people's data, including his own, is readily available on the dark web. Kayser emphasises the often complacent attitude many individuals possess, thinking they are not significant

enough to be targeted. Yet, as studies have shown, combining even seemingly innocuous data points can reveal unique identifiers: for instance, using just a ZIP code, gender, and date of birth can uniquely identify a vast majority of the U.S. population. In conclusion, humans themselves are the core issue, and there is an urgent need for individuals to be more protective of their personal data.

7.1.3 AI in Social engineering

The rapid evolution of AI poses a significant challenge in the realm of social engineering. C. Hadnagy's experience with ChatGPT-4, which crafted a convincing phishing email instantly, and the rising sophistication of voice generation tools underline this concern. Current defenses are ill-equipped to handle these AI-driven tactics. Organizations must proactively recognize and adapt to this reality. This means updating defense strategies, training employees to discern AI-generated content, and reassessing current cybersecurity strategies with AI-specific threats in mind. Immediate action will safeguard organizations against the growing AI-enhanced social engineering threats.

7.2 Conclusion

Several key insights emerged in response to the initial problem statement of understanding how organisations raise cyber awareness and fortify against social engineering attacks. Firstly, attackers utilise various social engineering techniques that significantly influence user behaviour and choices within organisations. Addressing the question of common countermeasures, both social engineering countermeasure tools and awareness training emerged as predominant defences. However, their effectiveness can be various and context-dependent. By looking into best practices, it became evident that organisations need a two-pronged strategy: raising acute awareness of social engineering threats and concurrently enhancing their overall cyber hygiene. This project highlights that answering the formulated problem requires a holistic view, blending technology and human-centric approaches.

Bibliography

[1] Linda Rosencrance. *Definition social engineering*. URL: https://www.techtarget.com/searchsecurity/definition/social-engineering.

[2] Kevin D. Mitnick. *The Art of Invisibility: The World's Most Famous Hacker Teaches You How to Be Safe in the Age of Big Brother and Big Data*. Little, Brown Company, 2019.

[3] Wenni Syafitri et al. "Social engineering attacks prevention: A systematic literature review". In: *IEEE Access* 10 (2022), pp. 39325–39343.

[4] Social-Engineeri.org. *Social engineering - Everyday People*. URL: https://www.social-engineer.org/framework/general-discussion/categories-social-engineers/everyday-people/.

[5] FIRST. *Common Vulnerability Scoring System version 3.1*. URL: https://www.first.org/cvss/v3-1/cvss-v31-specification_r1.pdf.

[6] PurpleSec. *Social Engineering Best Practices*. URL: Social%20Engineering%20Best%20Practices/.

[7] ProofPoint. *What Is Social Engineering?* URL: https://www.proofpoint.com/us/threat-reference/social-engineering.

[8] Roger Grimes. *GLBA and Other Regulations Wake Up to the Importance of Security Awareness Training With June 9, 2023 Deadline*. URL: https://blog.knowbe4.com/glba-and-other-regulations-wake-up-to-the-importance-of-security-awareness-training.

[9] World Economic Forum and Accenture. *Global Cybersecurity Outlook 2022*. Tech. rep. 2015.

[10] ProofPoint. *2022 State of the Phish*. Tech. rep. ProofPoint, 2022.

[11] jamf. *Phishing Trends Report 2021*. Tech. rep. jamf, 2021.

[12] Sikker Digital. *Sådan spotter du: Falske mails og sms'er*. URL: https://sikkerdigital.dk/borger/spot-svindel/saadan-spotter-du-falske-mails-og-smser.

[13] Centre for cyber security. *Cyber threat from phishing emails*. Tech. rep. CFCS, 2020.

[14] Fortra - Agari. *Ransomware Attacks: Why Email Is Still THE Most Common Delivery Method*. URL: https://www.phishlabs.com/blog/ransomware-attacks-why-email-is-still-the-most-common-delivery-method/.

[15] Thomas Brewster - Forbes Staff. *Fraudsters Cloned Company Director's Voice In $35 Million Bank Heist, Police Find*. URL: https://www.forbes.com/sites/thomasbrewster/2021/10/14/huge-bank-fraud-uses-deep-fake-voice-tech-to-steal-millions/?sh=21235f5f7559.

[16] Digitaliserinstyrelsen. *Pas på phishing: Falske SMS og e-mails om MitID i omløb*. URL: https://digst.dk/nyheder/nyhedsarkiv/2022/august/pas-paa-phishing-falske-sms-og-e-mails-om-mitid-i-omloeb/.

[17] Enisa1. *Phishing/Spear phishing*. URL: https://www.enisa.europa.eu/topics/incident-response/glossary/phishing-spear-phishing.

[18] Barracuda. *Spear Phishing: Top Threats and Trends*. Tech. rep. Barracuda, 2021.

[19] BBC. *Major US Twitter accounts hacked in Bitcoin scam*. URL: https://www.facebook.com/marketplace/item/7015258078491508/?ref=search&referral_code=null&referral_story_type=post&tracking=browse_serp%3Ab577e668-7599-44de-abad-78cdc04857ed.

[20] Enisa2. *What is Social Engineering?* URL: https://www.enisa.europa.eu/topics/incident-response/glossary/what-is-social-engineering.

[21] The Register. *Half of people plug in USB drives they find in the parking lot*. URL: https://www.theregister.com/2016/04/11/half_plug_in_found_drives/.

[22] Amin Azmoodeh et al. *Deep Fake Detection, Deterrence and Response: Challenges and Opportunities*. URL: https://www.researchgate.net/figure/Cyber-Kill-Chain-6_fig2_365820432.

[23] Crowdstrike. *WHAT IS THE CYBER KILL CHAIN? PROCESS MODEL*. URL: https://www.crowdstrike.com/cybersecurity-101/cyber-kill-chain/.

[24] TechCrunch - Carly Page. *North Korean hackers impersonated journalists to gather intel from academics and think tanks*. URL: https://techcrunch.com/2023/06/06/north-korea-hackers-kimsuky-strategic-intelligence/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmRrLw&guce_referrer_sig=AQAAAIizRTnKem_FJXFB7RNJEBe_nx1fmLHXUwNxiL530yOgoeN7DoFHixo3Tbo2Ovgi
UINf9sJWYTr6M69BhDxFe_cKnt9kDVujlgAePDFjiViCg60_kFgilS9j4Zspac4gVUeB3f.

[25] Forbes - Michelle Drolet (CEO of Towerhall). *Seven Tips For A Successful Security Awareness Training Program.* URL: https://www.forbes.com/sites/forbestechcouncil/2019/08/16/seven-tips-for-a-successful-security-awareness-training-program/.

[26] Fortinet - Training Institut. *2023 Security Awareness and Training.* URL: https://www.fortinet.com/content/dam/fortinet/assets/reports/report-2023-security-awareness-and-training.pdf.

[27] "Osterman Research". *"Cyber Workforce Resilience Trend Report"*. Tech. rep. " Osterman Research", 2023.

[28] Christopher Hadnagy. *Social Engineering: The Science of Human Hacking, 2nd Edition.* Wiley, 2018.

[29] Security Week - Kevin Townsend. *Security Awareness Training Isn't Working – How Can We Improve It?* URL: https://www.securityweek.com/security-awareness-training-isnt-working-how-can-we-improve-it/.

[30] Cisco. *What Is Social Engineering?* URL: https://www.cisco.com/c/en/us/products/security/what-is-social-engineering.html.

[31] BlueSteel Cybersecurity. *BYOD (Bring Your Own Device) Cybersecurity – Can The Two Get Along?* URL: https://bluesteelcyber.com/byod-bring-your-own-device-cybersecurity-can-the-two-get-along/.

[32] Ivana Vojinovic. *Save Your Data with These Empowering Password Statistics.* URL: https://dataprot.net/statistics/password-statistics/.

[33] Lawrence Abrams. *MFA Fatigue: Hackers' new favorite tactic in high-profile breaches.* URL: https://www.bleepingcomputer.com/news/security/mfa-fatigue-hackers-new-favorite-tactic-in-high-profile-breaches/.

[34] Rossouw Von Solms and Johan Van Niekerk. "From information security to cyber security". In: *computers & security* 38 (2013), pp. 97–102.

[35] Dr. Chuck Easttom. *Computer Security Fundamentals Fourth Edition.* Pearson IT Cybersecurity Curriculum (ITCC), 2019.

[36] Limin Sun Zuoguang Wang and Hongsong Zhu. "Defining Social Engineering in Cybersecurity". In: *IEEE Access* 8 (2020). DOI: 10.1109/ACCESS.2020.2992807.

[37] Joseph M Hatfield. "Social engineering in cybersecurity: The evolution of a concept". In: *Computers & Security* 73 (2018), pp. 102–113.

[38] Fatima Salahdine and Naima Kaabouch. "Social engineering attacks: A survey". In: *Future Internet* 11.4 (2019), p. 89.

[39] Katharina Krombholz et al. "Advanced social engineering attacks". In: *Journal of Information Security and Applications* 22 (2015). Special Issue on Security of Information and Networks, pp. 113–122. ISSN: 2214-2126. DOI: https://doi.org/10.1016/j.jisa.2014.09.005. URL: https://www.sciencedirect.com/science/article/pii/S2214212614001343.

[40] Charles McFarland Raj Samani. *Hacking the Human Operating System The role of social engineering within cybersecurity*. Tech. rep. Intel Security (McAfee), 2015.

[41] Maha Rita Arabia-Obedoza et al. "Social Engineering Attacks A Reconnaissance Synthesis Analysis". In: *2020 11th IEEE Annual Ubiquitous Computing, Electronics Mobile Communication Conference (UEMCON)*. 2020, pp. 0843–0848. DOI: 10.1109/UEMCON51285.2020.9298100.

[42] Dean F Sittig and Hardeep Singh. "A socio-technical approach to preventing, mitigating, and recovering from ransomware attacks". In: *Applied clinical informatics* 7.02 (2016), pp. 624–632.

[43] Liu Xiangyu, Li Qiuyang, and Sonali Chandel. "Social Engineering and Insider Threats". In: *2017 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)*. 2017, pp. 25–34. DOI: 10.1109/CyberC.2017.91.

[44] Mohammad Hijji and Gulzar Alam. "A multivocal literature review on growing social engineering based cyber-attacks/threats during the COVID-19 pandemic: challenges and prospective solutions". In: *Ieee Access* 9 (2021), pp. 7152–7169.

[45] Harjinder Singh Lallie et al. "Cyber security in the age of COVID-19: A timeline and analysis of cyber-crime and cyber-attacks during the pandemic". In: *Computers Security* 105 (2021), p. 102248. ISSN: 0167-4048. DOI: https://doi.org/10.1016/j.cose.2021.102248. URL: https://www.sciencedirect.com/science/article/pii/S0167404821000729.

[46] Raphael Hoheisel et al. "The development of phishing during the COVID-19 pandemic: An analysis of over 1100 targeted domains". In: *Computers Security* 128 (2023), p. 103158. ISSN: 0167-4048. DOI: https://doi.org/10.1016/j.cose.2023.103158. URL: https://www.sciencedirect.com/science/article/pii/S0167404823000688.

[47] Vahid Garousi, Michael Felderer, and Mika V Mäntylä. "The need for multivocal literature reviews in software engineering: complementing systematic literature reviews with grey literature". In: *Proceedings of the 20th international conference on evaluation and assessment in software engineering*. 2016, pp. 1–6.

[48] Robert B. Cialdini. *Influence, New and Expanded: The Psychology of Persuasion*. Harper Business, 2021.

[49] Zuoguang Wang, Hongsong Zhu, and Limin Sun. "Social Engineering in Cybersecurity: Effect Mechanisms, Human Vulnerabilities and Attack Methods". In: *IEEE Access* 9 (2021), pp. 11895–11910. DOI: 10.1109/ACCESS.2021.3051633.

[50] Markus Huber et al. "Towards Automating Social Engineering Using Social Networking Sites". In: *2009 International Conference on Computational Science and Engineering*. Vol. 3. 2009, pp. 117–124. DOI: 10.1109/CSE.2009.205.

[51] Matthew Edwards et al. "Panning for gold: Automatically analysing online social engineering attack surfaces". In: *Computers Security* 69 (2017). Security Data Science and Cyber Threat Management, pp. 18–34. ISSN: 0167-4048. DOI: https://doi.org/10.1016/j.cose.2016.12.013. URL: https://www.sciencedirect.com/science/article/pii/S0167404816301845.

[52] Hussain Aldawood and Geoffrey Skinner. "Educating and raising awareness on cyber security social engineering: A literature review". In: *2018 IEEE international conference on teaching, assessment, and learning for engineering (TALE)*. IEEE. 2018, pp. 62–68.

[53] Ryan Heartfield and George Loukas. "Detecting semantic social engineering attacks with the weakest link: Implementation and empirical evaluation of a human-as-a-security-sensor framework". In: *Computers Security* 76 (2018), pp. 101–127. ISSN: 0167-4048. DOI: https://doi.org/10.1016/j.cose.2018.02.020. URL: https://www.sciencedirect.com/science/article/pii/S0167404818301780.

[54] Ozgur Koray Sahingoz et al. "Machine learning based phishing detection from URLs". In: *Expert Systems with Applications* 117 (2019), pp. 345–357. ISSN: 0957-4174. DOI: https://doi.org/10.1016/j.eswa.2018.09.029. URL: https://www.sciencedirect.com/science/article/pii/S0957417418306067.

[55] Scibbr. *https://www.scribbr.com/methodology/qualitative-quantitative-research/*. URL: https://www.scribbr.com/methodology/qualitative-quantitative-research/.

[56] Latanya Sweeney. "Simple demographics often identify people uniquely". In: *Health (San Francisco)* 671 (2000), pp. 1–34.

Appendix A

Written Interview With Bank

Hac: Can you walk me through the process by which the bank conducts its assessments related to social engineering vulnerabilities, such as phishing attempts or impersonation attacks? What are the key focus areas and challenges in these assessments?

Bank: Such a process is multi facetted. We make use of threatmodelling in which case social engineering aspects are also assessed for the final solution. The key challenges in such assessments and decisions in terms of deciding security level is the balancing of Economy (cost of implementation), Usability (ease of use) and Security (restrictions, checks, tests etc)

Hac: What types of tools or technologies (e.g., email filters, behaviour analytics) does the bank employ specifically for countering social engineering threats? How do these integrate into the broader cybersecurity framework of the bank?

Bank: We do not comment or disclose information about security measures

Hac: On what basis does the bank select its methods and tools to combat social engineering? Are there any industry standards, benchmarks, or peer recommendations that influence these decisions?

Bank: Not quite sure I understand the nature of this question

Hac: How does an employee's previous IT experience or lack thereof shape the way you approach their training in social engineering countermeasures (e.g., different training modules for tech-savvy vs. non-tech-savvy employees)? Are there any challenges or advantages you've noticed in training individuals from diverse IT backgrounds?

Bank: Not quite sure I understand the nature of the question. We don't distinguish training efforts based on professional background and experience. We prepare

training based on objectives that stem from needs and market demands. Of course, coming from different educational and professional backgrounds provide different grounds for understanding all domains of IT security:

- Engineering backgrounds are typically strong in security architecture
- Data science majors are typically strong in application security and development
- Business administration majors and economists are typically strong in governance and risk management

Hac: How is the social engineering awareness training structured for different employees? Is the training customized based on roles, IT experience, or background (e.g., HR vs. IT staff)? If so, could you provide examples?

Bank: Training is compiled against various user groups in the organization. Developers receive a different set of objectives and requirements – more specialized towards security – whereas other users received "basic" training. So, yes, training is diversied dependent on target group and tasks/responsibilities.

Hac: In situations where employees have more flexibility in how they use their devices (e.g., bring your own device), such as working remotely, how do your strategies and policies for device management address social engineering threats like compromised devices or stolen credentials?

Bank: We run a very small footprint of BYOD, and where this exist the company has deployed security measures and fully control the device's security posture.

Hac: How does the bank measure the effectiveness of its countermeasures against social engineering (e.g., simulated phishing campaigns, employee surveys)? Is there a feedback loop in place to continually refine and improve upon existing methodologies?

Bank: There are a variety of data points that we consider, but a significant metric Is the volume of attempted fraud, successful fraud attempts, monetary volume of the attempts and many more which we do not disclose.

Hac: With the evolving nature of social engineering tactics, like deepfakes or vishing, how does the bank plan to adapt and modify its countermeasures in the coming years? Are there any emerging trends or threats in the field of social engineering that you're particularly watchful of?

Bank: Needless to say that AI has provided leverage for even more sophisticated fraud techniques and we are observing these trends heavily alongside investigating countermeasure.

Appendix B

Chris Hadangy

Hac: How about giving me a short intro of you? Who are you? What is your background?

C.Hadnagy: Sure. So Chris Hadnagy, I run a company called Social Engineer LLC. It started back in, I started the company like in 2009. I wrote my first book after writing the world's first framework on social engineering. I did that because I wanted to understand how the things I was doing during adversarial simulations were working. I would send phishing emails or make phone calls or bypass security and everything worked, and I didn't know why. So I started reading all these books you can see behind me on psychology and influence and persuasion, nonverbals and from that I developed a framework that would help everyone understand how communications can be used as an influence tool and how that influence can be used to test and audit security so since then, since 2010, I've written 5 books on the topic, I work with many of the world's leading scientists and researchers and trying to understand human decision making so that way we can learn how to fix the problems and presently I'm working with some AI experts to understand how that's gonna be used in social engineering coming up, so that's a little bit, Ohh I run a non-profit on the side that uses all these skills to help track people who traffic children and create child abuse material.

Hac: Hmm. I have listened to some of your podcasts and that did ring a bell. Just a light introduction into what we're talking about with social engineering, what is social engineering, and you know, specifically, maybe even to you. What is social engineering to you, and what are the misconceptions about social engineering?

C.Hadnagy: I love that question. So yeah, I like it because when you look up social engineering, the definition that most often pops up is the use of manipulation to get someone to do something, right. I don't really like that definition because I think that when you word it that way, you're saying that social engineering is always negative. So when I wrote my first book, I came up with a brand new definition, which is any act that influences a person to take an action that may or may not be in their best interest. I use a broad definition cause I think there are positives, right? Like we can use influence tactics or social engineering tactics to get someone to be healthier or to get someone to change a lifestyle habit that might be bad for them to eat better to exercise more, or you can use it to break in a building right? - or sending phishing e-mail so the same principles apply though, and I use these things as a parent. I have two children, I use these things all the time to get compliance from my kids and it's not manipulating them, it's using influence and good communication tactics to get people to do things so to me. The final part of your question, the misconceptions is that social engineering is always negative. That's a big misconception that you can use it in very positive ways and understanding how to use these skills can actually make you a better communicator in life.

Hac: Hmm, OK. You know, I actually took that, I think it took that from your book from the report I'm writing about the, you know how social engineering is actually applied in everyday life. Yeah. So that's an example too. It gives an angle to understanding, really, what social engineering really entails, rather than is just a mail. We'll get into that. I think before moving on to specific social engineering attacks, I want to ask you something that I read on your website. On your website description, you know socialengineer.com, that you use scientifically proven methodologies to uncover vulnerabilities and define risk and et cetera. How does this methods differentiate from other conventional methods?

C.Hadnagy: Yeah. So one of the things I'm really happy of is we have a scientist on board, a young woman who has her PhD and and she understands psychology, nonverbals, behavioural analysis. One of the things that we do is like, let's say for our vishing calls is we have defined principles of influence. We ask our vishers that when they're going to make a call to pick a particular type of influence that they're gonna use and then to use it and then document how did it work or did it work? Did it fail?You know, so that way when we go back to a client, we can say, hey, the reciprocation was used on this call and reciprocity

worked really well with your people, whereas authority didn't work too well. So let's enhance the education around gift giving and reciprocity. You're doing really good on authority and that helps us to pinpoint where weaknesses are, right. It's not enough to say ohh my back hurts, If you go to a doctor, my back hurts. Well, where is it? Upper back, lower back, middle back, the side. Is it a bone? Is it a muscle? Right. If you just say my back hurts, you can't really fix the problem. So if you just say ohh you're weak in social engineering. But where? What? What scientific principles are being used against you that are working, because a lot of companies have strength in some things and weaknesses and others, so we use those methods to pinpoint whereas what we find with a lot of other companies and this I'm not trying to say anything negative but as they just use social like they'll just fish or they'll just break in and it's kind of like being an auto mechanic that doesn't really understand, how cars work so to really fix the problem, right? - to do what your study is trying to prove, you have to understand how humans work. Otherwise, how can you fix the weakness? So that's why we use the scientific method.

Hac: I have another question which is social engineering attacks focused. I you know, looking at the evolution, you know, the development of social engineering attacks. Certain legacy attacks such as, you know, mentioned tailgating, you know, going to secure facilities or just facilities in general or dumpster diving for documents, these have become more challenging to to this heightened security and digital transitions. However, as some abilities close, another opens. So in your experience, how does you know social engineering techniques have evolved from what these more traditional to more current and the organisational landscape? Are there any old-school techniques that are in your opinion still a threat?

C.Hadnagy: Yeah. You know, I think what happens with new technology is, sometimes, when new technology comes out, we see a quick increase in security, right? So let's say MFA—multi-factor authentication—that came out, and it seemed like the fix. Like, man, we couldn't bypass it. Now, years go on, and we get used to it so much. It becomes something that is not new or novel. So now, we have about a 60% ratio on bypassing MFA, right? Just by calling people up and saying, 'Hey, I'm gonna send you—this is Paul from IT—I'm gonna send you a code because I need to verify your system is still in your control because we got some bad traffic.' And then, I'm on their website; I hit the button, they get the code, they read it back to me. I now have their MFA code, and I

can log in and change their password, and they're locked out of their account. And that works like 60% of the time.

You know, why? Why does it work that much? Well, because, you know, when MFA first came out and it was new, people were focused on it. But now, it's something that we use for everything. Everyone, in my accounts, right? My duo thing has pages of where I have to go get codes. Right. My password manager has pages of those codes because it's on every account I have. So, you look at one of those big hacks that happened last year where they did MFA flooding, right? Where the attackers just sent a non-stop barrage of requests. Eventually, the user got, like, 'Why are all these coming in?' and he hit the wrong button, and he hit 'allow,' and now the people got in. So, fatigue, overworking, getting tired, and when technology is no longer new—when it's not novel anymore—then we see the same kind of attack vectors working that have worked forever, right?

But with that said, I would say there are still times—and I can't believe it—that we go dumpster diving. I find unshredded documents. This just happened recently; we were on a job. Grab a garbage bag, open it up, and there's a whole bunch of financial statements, not shredded, right? So, when we report it, we're like, 'Can we find out why this wasn't shredded?' Well, the person spilled coffee on their desk. They thought in their mind, 'Well, the paper's wet; I can't shred it.' So they just took all the papers and put them in the garbage can. Right? So they did something that they thought, 'Well, no one's gonna. I mean, it's all ruined; it's all wet, covered in coffee.' Well, great; when we got it, it was dry. We dried it out. Now we had just some dirty financial records. Right?

The user didn't think about shredding them because of the situation. So, we still see insecure practices occurring because people get fatigued, get tired. You know, when the pandemic happened, and we all started working from home, that opened up a whole other layer of insecurity for us. Because now, we're sitting at our home computers, and our kids are running in and maybe using our computer. Or we're on our home network that does not have all the protocols and security that we have at work. Doesn't have the firewalls that we have, and now you have everybody from massive companies around the globe all connected to weak networks. And trying to do the same secure work. So, we saw a massive increase in problems because of that.

Hac: OK, so if you would have done, or you would have done a Google search

on what social engineering is, even if it's 5-year-old research papers, it seems there's still somewhat—not lacking in knowledge, but—stuck in that basic definition of what tailgating is, what reciprocity is, what is quid pro quo. So, when I was about to research about it, I really had to like really look for knowledge and not find the same knowledge all the time. So it seems, is this these definitions typically, -maybe there is some cases explained for each of the of the social engineering types of attacks, but it still isn't there more like social engineering 2.0?

C.Hadnagy: So one of the reasons I hired Dr. Abbey is so we could start performing research and putting out new information that would help with this problem. Because, I agree, when we talk about social engineering attacks, we're still looking at the four basic types: phishing; which is email-based vishing; which is voice phishing, smishing; which is SMS phishing, and then you have impersonation; which could be in person or through social media.

So all social engineering is encapsulated in those four areas. The challenge for people is that when you talk about social engineering, it's not like network attacks where, before, it was viruses, then it was Trojans, and now it's ransomware. These have advanced, and we see new types of attacks on networks and computers all the time. The four base attacks are the same, minus smishing, and they've been the same since the beginning of man. Think about it: we didn't have email, but there were scams happening in person, and letters were sent by con men. You can go back to the 1800s and read about Victor Lustwig and people like that, and they were conning people. So, these attacks are not new and they're not novel. What happens is the medium in which we perform them is what becomes new. I don't know about the country you're in, but if you go back four years, I hardly got any smishing attacks. Once in a while, it would be so lame, like some girl saying 'I met you in college, you wanna hang out?' But now, I'm probably getting 10 to 12 SMS messages a day that are bad. And that's because our cell phone numbers have been hacked and are floating around, leading to a constant barrage of attacks. So, the way that they are doing it is not new, but the method hasn't. The challenge when we write about it is that it feels like we're writing the same exact thing over and over again because the same types of attacks are happening and we haven't yet come up with a solid methodology to patch the holes and fix it.

Hac: OK. So you mentioned phishing, vishing, and smishing, and then impersonation. There seems to be a definition separating those three

phishing methods with impersonation, it seems, which also is demonstrated in my way of explaining these social engineering attacks. I separate them by saying there is an execution method, which is like more how you distribute your attacks, whether it's on the phone call, SMS, or by email, which are the most standard. Then there is an online version, which needs a little bit more elaboration because that can be done in several ways. But then there is impersonation, which I see, which I think has, which is more of a technique, isn't it? Because you can do impersonation via vishing, or smishing, or phishing.

C.Hadnagy: You can. Yeah, so you can. But when we talk about—so like again, I like your, I like the way you define it—when we talk about our medium of delivery, is it phone, is it text message, is it email, that's good. Impersonation could be delivered through all three of those. But it also can be a separate attack, in the sense like we see LinkedIn impersonation non-stop. Right, there was a big attack on U.S. military where an Iranian hacker group started LinkedIn accounts for young, attractive female reporters, and then they reached out to military commanders that were about to retire and asking to interview them. They interviewed them all through LinkedIn. It was an Iranian hacker group, it was not. None of them was real, and that was an impersonation attack where the military generals thought they were speaking to these reporters. They weren't; they were speaking to a hacktivist group. So we see that all the time. Or then it can be people making believe they're the IRS or that they're the government and you owe money or something to that effect. So the method of delivery for impersonation does alter. But impersonation could be in person, right? We see that all the time too, like people acting like they're a fellow employee walking into a company, stealing things. So we, you know, we see these different methods for impersonation and we can see combinations, right. I see a lot of phishing and vishing mixed. Or a lot of smishing and phishing mixed, like where somebody will send an email and then text and say, 'Hey, did you get that email I just sent?' So we see a lot of combination attacks to try to perform the attack.

Hac: OK, I actually want to touch upon the definition thing. Maybe we can talk about that in the end. And so, I want to dive into some discussions about countermeasures—how to combat social engineering. I actually talked with a bank here in Denmark. He was, I think he was the CISO, but he wasn't officially labeled as such. He was the IT security manager, so that sounds like a CISO, but it didn't say so. Maybe because it was written in Danish. To answer your previous question about four years

ago in Denmark, how the smishing thing was, it was mostly state-based messages for things like MFA. We have this state-funded way of logging into communal accounts and all of that. But it boomed during COVID. So, countermeasures. The bank I talked to said that when making a final decision on what tool to use—both non-technical and technical tools to choose—they have to balance economic cost, implementation, usability, and then, in the end, security. So, I want to ask you, what is your experience in how top organizations strategically choose their tools or countermeasures, especially considering these factors? Are there any benchmarks or industry standards?

C.Hadnagy: Sadly, no, and there's reasoning for that, right? It's hard to benchmark this because not every company, let's say not every bank, will have the same security setup. They won't have the same protocols or educational services. So the issue is, how do you benchmark things when there's no baseline? To set a baseline, you'd have to have one bank that you say is the best on earth, and then you rate every other bank against that bank. And no one's done that yet. So the issue that comes in is exactly what you said.

Companies know they need help, they need to secure against these types of attacks. So they say, 'OK, phishing. What do I need?' Well, you need good IDS, IPS, you need firewalls, antivirus, and you need to set up your mail server so you tell people when it's external, that you need filters that look for email that might be dangerous, and then you need an education program to help people realize what to do to report phishing properly. You take all of that, and you need a giant budget, and this is just for phishing. We haven't even talked about vishing or smishing or anything else. So they look at this and they say, 'OK, 90% of all breaches have a phishing element to it. Let's spend most of our money there.' And sadly, this is just the nature of the world, the bad guys aren't limited by spending. If it takes them a year to infiltrate your network, they'll spend the year to do it. But our good guys, they're told, 'Here, you have $500,000. Spend it wisely.' And they've got to figure out how to spread all that money across all of their security needs. And that's a difficult process, especially when you have a million vendors telling you they're the best. And when you have a guy or a gal in the position who may not know everything about it, how do you know where to spend the money properly? So it's set up kind of like, I hate to say it, but it's like set up for failure, you know.

Hac: And in addition to that question, I think something that is obviously

interesting for my thesis is, you touched upon it a bit because you said in a way that it's every man for themselves. It's every bank is like, 'You need to develop your own thing,' which is not unreasonable, per se, but one might ask, is it reasonable? And as an aspect of that dilemma, I don't have the answer for that, obviously, but so doesn't that reflect the individuals working at that company? Do the employees have an effect on these decisions made in companies, in terms of which tool they're using, like their skills and abilities and all of this?

C.Hadnagy: So, I mean, yes and no, right? I think the answer to your question is it's too cyclical. The answer is, well, if you do a good job at creating a really effective cybersecurity awareness program, it will make your employees better at their jobs, which means you may not have to spend as much money on the tools, right? But someone's got to start somewhere for that cycle to get kicked in. Right? And it's not. I'll tell you, we've been working with a company, I think it's going on six years, and we've been doing vishing. And we do over 1,000 vishing calls a month for them. The first two or three years, there was very little adjustment. You know, we were winning and they were not. We kept saying, 'just keep doing the education, you're going to see it.' Now year four kicked in, and all of a sudden there was a big spike where we were failing and they were winning. They asked, 'what happened?' and we said, 'what happened is your people started getting the calls, learning the process, and reporting it properly.' Now, going on year six, we have a hard time as attackers, and that's good. They had to commit to that process for four years before they saw the change that made them realize there was ROI. So I look at a company like that and give them a lot of praise because they committed, you knew it was going to take a bit, but they committed to it and they have payoff because of it.

Hac: OK. More about the individuals in companies. You know, I can imagine, as you said, you've talked with numerous professionals. When it comes to awareness training and general policy making in the firm, security-wise, how do the different educational backgrounds influence the awareness training specifically? Are there any backgrounds more inclined to grasp it?

C.Hadnagy: Yeah, that's a really nice question. So this may not be a direct answer, but this made me think of something that I think is helpful. I've been seeing a lot more companies hire people who have degrees in psychology to become part of the cyber security awareness teams. I think that's smart because what you have is an IT guy, right? Let's say an IT

guy who gets it, but doesn't really talk well to others and he's running your security awareness program. Then he does things like we've seen it. Just last year, you know, GoDaddy, two years ago, GoDaddy makes an announcement to their company that there's no money for bonuses. The pandemic has hurt them, so no one's getting a bonus. The IT guy decides to send the phishing email out telling everyone, 'Hey, there's some bonus money. Click here to get yours.' I mean, that's horrific if you think about it.

Hac: Do you mean internally?

C.Hadnagy: Internally, he sent a phishing email to the employees as a test about bonuses. Now think about this: there's some single mom working her off at GoDaddy to make ends meet. And that bonus is the way she buys little Johnny's Christmas presents for December. Now she gets told there is no bonus, and she's worried. 'How am I gonna make Johnny's Christmas good? How am I gonna buy him presents?' And now she gets a phishing email that gives her hope. So she clicks the link and then she gets told, 'No, you were just a sucker and you have to go for training.' Now, that is horrible. So you find a company that hires a psychologist, and those things have to get run through. The psychologist is going to look and say, 'Whoa, that's going to mentally damage people.' So what I've seen is that companies who are spending more time with their cyber security programs, with psychology in mind, are actually doing a better job and seeing a positive increase in their companies.

Hac: Okay, so that makes sense for bigger companies to have that kind of element into it.

C.Hadnagy: And that's a hard one, right? Because you just said, 'So, how does a small company afford that? How does a small company go and hire?' Maybe you can't, but what you can do is what we try. We try to put out a lot of information. So, I have someone who has a PhD in psychology. We ask her to help us write articles and education. Small companies can come to our website and get that stuff for free. They can take our blogs, newsletters, and podcasts, and use all of that to help develop programs based on psychological principles. We try our hardest to make our education available to the public so people can use it. Is it the same as having one on staff? No, but it's at least something. It's a start, so you're not just being hopeful.

Hac: You mentioned the reporting being an important element in actually trying to learn. Hey, when going in loops and trying to be better, especially if you know it takes four to six years. So how crucial is feedback looping and refining measures? Are there any practices you recommend for organizations continually evolving their strategies?

C.Hadnagy: Yeah, this is a great question. So think about this, this is my personal issue with a lot of phishing SaaS', software as a service. They focus only on click ratio, but click ratio is useless by itself. Let's say I'm an Amazon junkie and you're not, you never shop on Amazon. So you and I work for the same company and a phishing email comes out, it's an Amazon phishing email. I click it, you don't. Does that mean you're better at catching phishing emails than me? It doesn't. It just means that that theme didn't interest you. So if you send me a spa treatment phishing email, you know, go get my nails done, I'm not interested. So I'm not gonna click it, right? Whereas my wife, she'd click that in a second. So click ratio by itself is a useless statistic because it doesn't show the company anything. But reporting ratio, that tells you if your education is making a dent. So now what I say is that for phishing you grab four stats: those who clicked and didn't report, those are your worst employees. They clicked the email and didn't do anything to save the company. Those who clicked and did report, okay they did something bad, but then they realized, 'Oh, I shouldn't have clicked that,' and they reported it. That's good. Those who did not click and did not report, those people didn't do anything. Maybe they're not reading their email, it goes right to junk, they just deleted it. That's not good. You want to change that behavior. And those who did not click and did report, those are your star employees. Those are people who caught the phishing email and they reported it. When you see those numbers adjust in the right way, now you know that your program is actually being effective. That's how you prove ROI to the higher-ups, to the people who are writing the checks. Because they need to see that those ratios of people who are actually catching the phishing emails, thinking critically, and doing the right thing, that those are always going up. And to me, we don't focus enough on that in our education. We focus on getting really low click ratios or getting really low success ratios. And that's just one side of the coin. We need to focus so much more on reporting and we'll see a big difference if we do that.

Hac: How about, I just thought of something while you were explaining. When to report? So, the way you're describing it, it seems like, are we

in a simulation? Is it only a simulation? Because how would we know that they are not reporting something if it's not in simulation?

C.Hadnagy: Yeah, so two things. Yes, we want simulations reported, but how do we know? I'll give you an example. I worked with a company for six years. We did phishing tests for them and they have 250,000 employees. We phished all of their employees every month. After three years, they saw a 57% reduction in actual malware on their network. They started to analyze and realized it was because people started reporting actual phishing emails to the IT department. So it wasn't just our simulation phishing emails; they started reporting real phishing emails instead of clicking them. Those real phishing emails contained malware. So the bank, it was a big bank, was able to now capture that malware and stop it before it attacked the network. Or if someone did click it, they were able to stop the malware from spreading through the network. After five years, they had a 79% reduction in malware on their network. So it's not just simulated phishing emails, but simulated phishing helps us with stats that help us with ROI. So I can tell a company I'm working with, 'This is why you're paying us. This is what we're doing for you. This is why this is important.' And those simulated phishing stats help us with that. But we want them to report both real and simulated because the real ones are where they're going to start to see the actual effect in safety.

Hac: OK. So it's kind of ironic in some way because I would guess that these many, many employees knew they were being tested.

C.Hadnagy: So when we work with them, we always tell a company, 'Don't notify them every month, but tell employees, by the way, we run an internal phishing program and we're going to be phishing you. When you get a phishing email, here's what we want you to do.' And we tell them we want you to report it, and this is the method for reporting. We feel that's a better methodology for fostering a team-type atmosphere and cooperation. It's like if you went into a boxing gym and said, 'I want to learn how to fight.' If the trainer just wailed off and punched you in the face, you'd be like, 'Whoa, that's not cool.' But if he said, 'OK, put your hands up, I'm gonna show you how to block,' now you're ready, you're being prepared. So I think when we communicate correctly with our employees, we get better buy-in on the program and they actually want to be a part of it

Hac: Okay, would it be an unfair statement to say that it's not so much them being better at detecting phishing emails, but more about being

aware that it might be a test, and then actually catching real ones and stopping malware?

C.Hadnagy: So it's like any kind of muscle memory, right? If you've ever taken a martial art or boxing class, you stand in front of the bag and you do this. You do it for 10 minutes at a time and you're wondering, 'Why isn't this teaching me anything?' Then one day, you start doing combinations and it's not hard anymore. When you're in the ring with your sparring partner and he goes to throw a punch, you block. You think, 'Wow, I just did that.' That feels really good. So when you test people constantly with real things that are happening in the real world, they start to recognize, 'Oh, wait, I know what that is. That's phishing.' It gives them the muscle memory to actually defend properly. If they don't even know that this kind of attack exists, how can they defend? So half of awareness is getting them to realize, 'Hey, this is happening. This is what it looks like. This is what it feels like.' And you do it in an environment where you're not using shame or fear, but positive reinforcement. And those things help make an effective change for people.

Hac: What I was trying to convey is making the employees care. Like, make them feel they are at risk personally.

C.Hadnagy: Yeah. Yeah. So the way to do that is you have to make your education not just about your company. You have to make it personal. You're gonna tell them. Hey, guys. Look, right now you know this time of the year, there's a lot of attacks going against our kids. So let me tell you about some phishing that might be happening against your kids. Take that home and mourn your kids. When you take education and you make it personal for them. Like, here's how you know what right now in America, there's a lot of grandparent scams going on. And that's real. Like, they're calling grandparents up and stealing money from them. So you tell your employees, by the way, this month, there's a lot of grandparent scams. So we want to tell you how you can go home and keep your grandma. Your grandpa, your mom, your dad's safe when you make the education also personal and they start applying security principles in their everyday life. That will make them better, more secure employees. So it can't just be about work. You have to be willing to also spend some of your time and money educating them about personal things and that will make a bigger effect on them.

Hac: Hmm, okay. So, last question about the employees. I think we already discussed it, but maybe there's something to add. When it comes to

predefined cyber hygiene knowledge, how influential is the employees' general cyber hygiene or feedback? How influential is this on an organization when selecting? I feel like I'm repeating myself, but the rephrase of this question is: When do you know that you need to buy a feature into the company, like an extra layer of security? Is that something you don't do as much in IT firms or?

C.Hadnagy: No, it's a good question. I understand what you're asking. So I'll give you an analogy. I have a trainer for the gym. And when I started with him, he wanted to see where I was weak. So we did a lot of exercises that were like one-offs. Like, 'let's bench,' 'let's do squats,' 'let's do this.' And then he could see, okay, your form is good here, but your form's really bad there. So we need to work on that. Companies need to take the testing that we're talking about, phishing, vishing, and look at where they're weak. Then decide, is there some technology that can help our people stay safe, that can make their job easier? Or do we need to beef up education? So my trainer, he says, 'OK, go get wrist straps. That will help you when you're doing deadlifts because your grip is weak.' So I get wrist straps and that helps me accomplish the lift easier. Because he noticed a weakness and he gave me a fix. So from a company standpoint, instead of just blaming the employee and saying, 'oh man, look, you're so weak at this, fire them and get another person' look at ways that you can enhance the ability for your employees to fight against this. You'll see a difference. You really will. You'll see a difference in how they feel like you actually care about their security. And that will make them be more invested in your company.

Hac: OK. And the next question was about actually the effectiveness of evaluating metrics, and that's something you already touched upon. But let me ask the question anyway, so. You know, apart from the number of fraud attempts or, you know, how much cost, monetary value you have lost? What other metrics have you come across that organizations use for, you know, measuring the effectiveness of countermeasures? And you know, I defined these indicators or factors which we should look beyond, which is, uh, the number of successful scams, like how many there are, and how many incidents there have been, what the cost is. Is there anything else that we can relate to when you're adjusting the awareness training?

C.Hadnagy: Yeah, it's a hard question because there's no general answer, right? But I'm gonna refer back to that bank I was telling you about. Um, that bank looked at actual malware on their network. They knew

that every year they had 'X' amount of malware. When they saw a reduction, that was a statistic that helped them see effectiveness, right? They could tie that malware directly to phishing. Now, people were reporting more phishing, and that was reducing malware. So, for a company to be able to do that, they first need to look at what attacks are happening against them and what's working. Right? So, um, for example, my company, we get way more phishing than we do vishing, way more. We get way more smishing than we do vishing because my company's remote. We all work from our homes, and we all use our personal cell phones to work. So, for my company, as small as we are, we look at, 'Okay, well, phishing and smishing are the two attack vectors that affect us.' So what can we do to help keep our company safe and our people safe from those two things? Right. It doesn't mean we ignore vishing. It just means that we're not focused on that as much because it's not the vector that's affecting us. Right, so companies to be able to do that, they have to understand where their weaknesses are. Evaluate how you can test those weaknesses. So what we do is we run phishing campaigns. Again, I only have 15 employees, but we run phishing campaigns against them every month, me included. Even as their CEO, I get phished every month, right? And it's important that you have that from the top down. So, I don't send the phish; one of my other people does, so that way, I can be phished. Because I have the most access in the company. If I fall victim to a phish, the company is gonna get ruined, right? So I should definitely be getting tested. If I fail, then I get training just like any other employee, right? So, that kind of top-down approach is something that we have to have in companies. You can't say, 'Oh, this is for everyone else but not me. I'm the CEO. I don't get—' Nope, that's really bad methodology because if you're not testing everybody, then the person you're not testing, that's probably your weak spot.

Hac: In some cases, actually, it is best to catch like the one on the top because they have the authority to do more.

C.Hadnagy: Yeah, 100% right. I have access to the bank account. 99% of my employees don't. I do, like my COO does, and my controller does. So why are they gonna? - They're not going to try to phish my social media person. She's got no access to the bank account, right? So they can try all they want, but she can't even do a wire transfer. So you wanna attack our company and you want money? You're going to go to the people who have the money and have access to the money, right? So yes, I should definitely be getting tested because I have the most

access to hurt the company. 100%.

Hac: There has been some incidents where there has happened where the manager of some kind has been tricked using deep fake on the phone and then something via a hacked email. You know, the phone call plus the sent email. Then like millions of dollars were sent because they were able to. But it was just a normal employee probably wouldn't be happening.

C.Hadnagy: 100%, right. So, um, you know, Toyota lost $34 million to three wire transfers because they contacted the person who had authority to do the wire transfers, right? So attackers are smart. We have to be as smart as them in preparing our training and our protections. Otherwise, you know, they're gonna. Like I said, they're not limited by time, money, effort, desire; they have it all. So they're gonna do what they need to do.

Hac: I don't want to spend too much time on the next question, but it's something that comes from my own heart. I've been trying to reach companies and initially had high hopes of conducting experiments with them. I was surprised that many organizations understandably refrain from discussing specific security measures publicly or even just with me. There might be concerns about exposing operational security, potential liabilities, and reputational concerns. However, the objective of my research is not to expose vulnerabilities, but to understand and highlight the best practices, right? So, from your perspective, personal or otherwise, what insights do you might have for researchers like me to better navigate these challenges in order to have better dialogue about social engineering counter measure?

C.Hadnagy: It's hard, you know. I had another guy who was doing some research, and what I did to help him is, you know, we got together on LinkedIn and I posted to companies saying, 'Hey, if anybody is willing to help by answering some questions for this guy,' and he got some help. I could do something like that to help you, but the hard part is, that a lot of companies are very skittish about sharing too much information. Even though your intentions are good, they don't know that. So they don't want to put information out there and all of a sudden see their name in some report saying, 'Look, this company's not doing enough,' or 'They're not doing good.' So they'd rather just sit there quiet. I don't have a great answer for that because I have the same problem. It's not easy.

Hac: The bank I was speaking with was very open and then we got closer to the deadline. They were very hesitant. Then I said in the end, let's just how about I just send you some question and you can reply me written and there was very diplomatic and formal about the answer. So it was..

C.Hadnagy: You might have a better time finding companies like mine that work with a lot of different companies and could give you information on things that are being done, but without giving you client names, right? So like, being able to describe things. Now, the thing that will be different is, I won't have intense knowledge of everything. Like, I know what I did for that bank I told you about; I know what services I performed for them and how they did. But I couldn't tell you in-depth everything they did for security. So that might make it biased or not provide enough detail for what you need.

Hac: It is, you're right. It's a difficult question to answer. It's just something that I wanted to share with you; that's why I was a bit persistent with you

C.Hadnagy: Yeah, which is OK. I need that.

Hac: Something for the next question. Something you've already touched upon is 'bring your own device,' and you know, in the same context—not necessarily—but just to make it into one question: remote work. And, you know, bring your own device and the increase of remote work have added, like you said, another layer of complexity to general security, not only social engineering. How have organizations been able to adapt to this path which social engineering has opened, and...

C.Hadnagy: Yeah, a lot of caution, right? So for us, like for someone to be able to use their personal phone for corporate email, they have to allow us to install an app. We can't see anything on their phone; we can't see their pictures; we can't access their phone. But if they were to quit or get fired, we have the ability to remove all corporate access from their phone. So we hit a button, and it deletes all the email and their corporate access. If they don't agree to that, then they can't have email on their phone, right? So, and then from remote work, they're not allowed to use a personal machine. You have to use a machine we give you, right? And we tell them, for a laptop, don't put anything personal on it because we're backing it up, we're monitoring it. This machine is just for work, and it has access to our VPN. So when you

need to get to our servers, you use our VPN. We don't allow anyone to use any personal devices for computers. If you want to use a personal device just for email, you have to give us the ability to install. So, I think a lot of companies are going that route where they're allowing remote work, but you have to have the same technological setup. If you're security conscious, you will. If you're not, you may not do that, and that's not great.

Hac: So it is a grey area, no own devices.

C.Hadnagy: It's, yeah. For our company and for a lot of companies we work with, they have the same rules. You can't use a personal device to do work. You can only use your personal device for email. But if you're accessing our servers or our data, it has to be on a machine that we own. They can try, but they couldn't even gain access to ours because the certificates for the VPNs are installed on the machine we give them. So they can't gain access to our data or servers from a personal machine. We have things locked down so that the way they gain access to stuff is very limited

Hac: What about using the browser - You can use the browser for many-many things.

C.Hadnagy: Yeah, so we tell people not to use our computer for anything personal, and people do. But what I tell them is, if you log into your bank account on our machine, that machine's getting backed up every five minutes throughout the day. We now have a backup of your passwords. Right? So, I'm not going to do anything with it. I'm not going to access your bank account.

Hac: And it's the app that does that or..?

C.Hadnagy: Yeah, so we have apps installed on our laptops that we use that have automatic backups. We're part of the Microsoft Cloud, so everything is automatically backed up to SharePoint. We have remote control software, so if an employee is getting fired or quitting, we can shut their laptop down and lock it from anywhere in the world. We have to do all those things because I have client data on those machines. If someone's going rogue, or quitting, or doing something wrong, I have to be able to control that. Otherwise, we're going to experience a breach. As a small company, a breach can shut us down.

Hac: I actually like that. It doesn't have to be some stranger hacking you. It's your own company. When you're exiting or something and you're about to do something bad, your company can hack you.

C.Hadnagy: Yeah, and that's why we say it has to be on our machine, right, I don't want..

Hac: Is that is it normal to have that access?

C.Hadnagy: In this country, I believe it is. I don't want to have access to your computer. Whatever you do in your personal time, whether you put your nude photos on there or watch a TV show, that's between you and you. I don't need to know that. So the machine that we give you is just for work. Just do work on it. You should have your own personal laptop for all other activities. We tell them that your kids, spouse, or partner shouldn't be touching this machine. Your dog shouldn't be touching this machine. This machine is just for you. We provide them with a laptop, keyboard, mouse, extra monitor, and headset. This is your office. Use this for work, and if you ever leave, we'll take it all back. That way, it separates their personal life from their work life.

Hac: Before we go into the last question, you mentioned something, and something else came to mind. What about the other way around, where you use your personal device for business?

C.Hadnagy: So the only thing that we allow is if they have their phone and they want to use their phone to access email. Then we allow that, but they have to install an app on the phone that gives our administrator access to delete that data off the phone if needed. We can't see anything. It actually creates a secure package on their phone. In that area, it's called 'Work,' as opposed to 'Home.' On my phone, I have 'Work' and 'Home,' and 'Work' is just my email and access to SharePoint. So if, let's say, I was getting fired, then the sysadmin would go in and hit a button that would delete my access to the corporate infrastructure and remove all of the email that was from my work on the phone.

Hac: No, I understand. Typically, when you're getting hardware from your company, you don't have any say on it. But it's another thing when you're using your own personal, but as far as I understand it, it goes both ways.

C.Hadnagy: Yes, but we don't allow any personal devices outside of that. So they can't use their personal laptop or desktop. Let's say we're a Windows

shop, and they love Mac. That's not my problem. When you come work for us, you're using Windows. I've had people say they love Mac, but that's irrelevant because we're not buying you an Apple device. Our software, backup, control software, SharePoint, everything we do is Microsoft-based. I have friends who work at other companies that are Apple shops, and everyone gets a MacBook when they get hired. My wife loves Apple and has a MacBook as her personal device, but she works for us and her work machine is a Windows device. She complained at first, but there was nothing I could do. This is how the company operates. Bigger companies can have blended networks with both Apple and Windows, but we're too small for that. Our control software works best on Windows, so that's why we are a Windows shop. It allows us the most control over everything. If we ever were to become a blended shop and somebody wanted to use a MacBook, we would have to buy it; you can't use your own laptop.

Hac: Okay, last question. It's about the future of AI in social engineering attacks, there's no doubt that in the future of social engineering attacks there's some element of AI technology. I'd love to hear your opinion. So, regardless, looking at the evolution of social engineering attacks over, let's say, the last 20 years, how do you foresee the progression of social engineering threats, like deepfakes or advancements in general AI tech? And given those threats, how do you foresee the progression of social engineering defense countermeasures? How do we adjust to the current countermeasures that we have? Are we even ready for it?

C.Hadnagy: I'll answer the last part first: we're not ready. The advancements in AI have happened so fast, and there's no regulation; no one's regulating it. They're pumping stuff out. The other day, I was messing with ChatGPT-4. Because I'm a public figure, I said, "Hey, you know, tell me about Chris Hadnagy." So it told me about me. Then I said, "Write me an email that will get Chris to come to my Russian dating site." It wrote a really beautiful phishing email. right? Because you can't say "phish" because it won't write a phishing email, but I said write me a marketing email that will get Chris to click on my Russian dating site. Beautiful, beautiful. Oh, that was wonderful, right? I'm thinking, "Crap, that was really good," and it did it in seconds. I can take an email that I'm writing, put it in chat, and say, "Hey, proofread this," and it corrects the grammar and spelling. So now you have people from foreign countries who are phishing in English, and they can have perfect grammar and spelling. We are quickly advancing to where voice AI is becoming more and more advanced. There was an attack here

in the States, guys downloaded this program Microsoft put out. With 15 seconds of audio, it can create a whole conversation in someone's voice. So they found some girl on Instagram, downloaded her videos, found her mother's phone number, called the mother, and said, "We have your daughter and we're gonna hurt her." They had her voice in the background saying, "Mommy, help me, help me." They demanded $50,000, and she said, "I don't have $50,000, I have $20,000." They took $20,000. They hang up; she does the transfer. Twenty minutes later, the girl walks in and she's like, "What the hell?" It was a scam.

So what's going to happen AI? - I don't know. The third part of your question was, "What do we do?" I think we have to start using AI to build defenses against AI. We have to start doing it. That's what I was saying before. We're right now working with a number of researchers to try to do that because we're not ready. I've been doing this for 20 years, and I'm not skilled enough to tell you if that voice is real or not. They now have an AI bot that can create a doppelganger of you. If you don't want to go to Zoom meetings, you can send it, and it will sit there and interact, and they won't know that it's not you. You train it, it may take a couple months. You giving it your voice and it watching you and watching videos of you and you gotta take lots of video and feed it into the bot and then it becomes you. Crazy, right? This woman is making millions of dollars right now. She she created a AI doppelganger of her and made a porn site and her fans can come and interact with the AI. So they could pay the bot to take the clothes off, and it's her. She took video of herself, and the bot now. She said she doesn't have to take her own clothes off now; she has AI stripping for her. Holy crap, right? Like, where is this going? AI is what, like, 5 years old? And we're here already. I mean, what's the next thing? I don't know, it's scary.

Hac: It just seems it's like a zero-day vulnerability. It feels like a zero-day vulnerability. If we cannot detect if a picture is fake, if the voice is fake or the video is fake, right? It seems like if we cannot do that, how should we recognize AI in other places than those three places?

C.Hadnagy: I don't have an answer. That's the scary part, I don't. Right? And even AI experts are saying it's getting to the point where they don't know how to tell you. Like two months ago, someone said, "Well, OK, but when an AI video is made, there's this glitch and you can see it, the pixels, this or that." They fixed that. That's what I mean, the advancements and what's happening with AI are happening so

fast, at lightning speed. We're not ready. We're just not. There's no preparations to defend, right?. Like I know it's a movie, but this is like how Skynet started, right? Because everything got built and then it became autonomous. It started to think for itself. I mean, I'm using AI right now in my non-profit to help with OSINT. It's unbelievably scary good. That's, that's horrifying to me. Like being able to feed it information and it finds connections that would take a human dozens of hours to find.

Hac: There were some researchers who actually did almost that, which was just writing a lot people by using a script that do crawling for information on LinkedIn and then messaging them. They actually got some good responses to it. Surprisingly, I think because they were not, they were not writing themselves but the bot did with what it was about. That was actually my last question. I think we can, you know, talk about AI for a long time. Thank you, Chris, for this enlightening conversation. It's been great discussing social engineering with you, I really appriciate it.

C.Hadnagy: that's OK cause I have a more meeting, so I'm glad. But, absolutely, I hope I have been helpful yeah. But um, I'd be more than happy to meet again, or try if you need more time.

Hac: You've definitely been helpful, Chris. Thank you for being with me. Take care and happy lunch or I don't know what time it is now.

C.Hadnagy: Yes. Yeah, it's it's 1:00 PM. So it's just about that time. It was nice meeting you.

Hac: You too. Goodbye.

C.Hadnagy: Bye bye.