Social Engineering in the Context of Cialdini's Psychology of Persuasion and Personality Traits

Bachelor Thesis

Susanne Quiel

19.07.2013

Supervisors:

Prof. Dr. Dieter Gollmann

Dipl.-Math.oec. Sven Übelacker

Hamburg University of Technology
Security in Distributed Applications
`https://www.sva.tuhh.de/`
Harburger Schloßstraße 20
21079 Hamburg
Germany

For my husband Niko, who has encouraged me to study Computational Informatics, and has supported me all the time. I love you.

Declaration

I, Susanne Quiel, solemnly declare that I have written this bachelor thesis independently, and that I have not made use of any aid other than those acknowleged in this bachelor thesis. Neither this bachelor thesis, nor any other similar work, has been previously submitted to any examination board.

Hamburg, 19.07.2013

Susanne Quiel

Abstract

This thesis shows that social engineering mainly relies on peripheral route persuasion and that consequently, Cialdini's principles of influence can be used to explain how social engineering attacks work. It is further shown by a comprehensive literature review that individual values of personality traits relate to social engineering susceptibility. Based on these arguments, a framework is proposed, which can help to guide future research. Suggestions to plausible relations between the personality traits of the Big 5 Theory and the principles of influence are made. These relations need to be evaluated by future research. Lastly, an integrated approach to prevention against social engineering attacks is proposed, which combines penetration tests using social engineering, security awareness trainings, and the development of a security-aware organizational culture. The methodology used in conducting this investigation is a theoretical, positive research approach. Extensive literature reviews have been conducted on social engineering, psychology of persuasion and the influence of personality traits on the success of social engineering attacks. Based on an analysis and discussion of these reviews, proposals for a new framework and a prevention approach have been developed.

Contents

1 Introduction

In late 2009, Robin Sage, a 25-year-old American cyber threat analyst, became active on multiple social networking websites. During a month, she made hundreds of new friends and connections, many of them working at American government entities like NSA (National Security Agency), DOD (Department of Defense), and Military Intelligence Groups as well as at Global 500 companies [59]. She received multiple job offers and gifts, and was asked to speak at security conferences. Her new online friends even asked her to review papers and presentations, and gave her access to email addresses and bank accounts. The drawback was that Robin Sage was not a real person. She was created by Thomas Ryan, a security specialist, as a social experiment about information leakage. Had Ryan been a criminal trying to use the information gained by this false identity, he could have stolen research, gained insight into some companies' goals, security, and salaries, and probably would have been able to commit identity and financial fraud.

This is a vivid example of social engineering (SE), which is broadly defined as a set of techniques used to manipulate people into performing actions or disclosing confidential information [48]. According to the Computer Security Institute (CSI) [58], exploitation of users' social network profile options were experienced by 5% of their respondents in 2010. 11% had their systems penetrated by outsiders and 39% had been fraudulently represented as senders of phishing messages, an "email based deception where a perpetrator (phisher) camouflages emails to appear as a legitimate request for personal and sensitive information" [72]. Another report states that "the most common attack vectors for social engineering attacks were phishing emails, which accounted for 47% of incidents, followed by social networking sites at 39%. New employees are the most susceptible to social engineering, followed by contractors (44%), executive assistants (38%), human resources (33%), business leaders (32%), and IT personnel (23%). [...] According to figures from the U.S. Computer Emergency Readiness Team (US-CERT), phishing attacks accounted for 53% of all security incidents in 2010 " [15].

The Verizon Data Breach Investigations Report [71] analysed a sample of 855 incidents in 2011, which occurred to organizations in 36 countries. Approximately 174 million records were compromised. 60 incidents occurred in larger organizations (with at least 1,000 employees), 660 in organizations with less than 1,000 employees. The organizations' size of the remaining incidents was unknown. These numbers show that every organization, regardless of size and geographical location, is a potential target for security attacks of all kind. External agents were responsible for 98% of data breaches, and 7% of the attackers adopted social engineering tactics. In larger organizations, social engineering tactics were used in 22% of the cases, indicating that the probability to be a target for social engineering

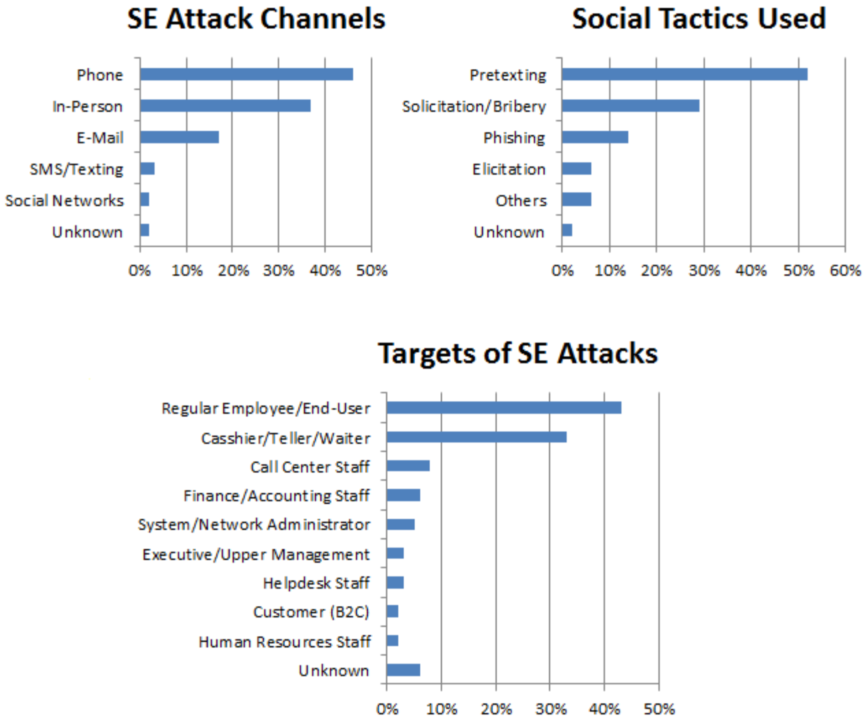

Figure 1.1: Social attacks, attack channels, and targets, Verizon Data Breach Investigations Report [71]

is higher for larger organizations. The authors of the report argue that this is probably due to greater anonymity of employees and more sophisticated perimeter defences. The third most used attack vector in large organizations was an external using social engineering attacks against people's integrity, while the eighth most used attacks went against confidentiality [71].

In Figure 1.1, the social tactics and channels used as well as the targets of the attacks are shown. The different kinds of attacks will be discussed in Chapter 2. For now, it should illustrate the point that social engineering encompasses a variety of tactics, which are often combined. The targets of social attacks enfolded people of all hierarchy levels in organizations. Regarding the channels used, it is likely that social networking sites are under-represented in this sample. It can be assumed that social networks are used for information gathering before the actual social engineering attack, since the information on social networks are not only easily obtained, their fetch is usually not detected, at least on public profiles. The detection of the use of this channel in a social engineering attack is very difficult since organizations cannot and should not control everything that the people inside the organization write on social networking sites. The organizations can of course use non-disclosure agreements, and prohibit private use of computers at work for preventing internal documents to be spread into the public domain. This could be enforced by tapping the outgoing data links of the organization. However, it is neither possible nor legal to control what employees write on social networks in their free time.

The statistics quoted in the previous paragraphs show that social engineering is on the rise, and threatening not only companies and government agencies, but also individuals, the latter mostly regarding identity fraud. Any person could be the target of such an attack. Of course, there might be differences in how great an asset a person is. A manager could possibly disclose more valuable information than a cleaning worker. However, one single person in an organization who gives an attacker information, no matter how insignificant it may appear, can be sufficient to enable him to compromise whole systems. In this regard, it is highly interesting to examine which factors contribute to the success of these attacks. The authors of the Verizon report state as a reason for the use and success of social engineering tactics that "the 'carbon layer' of information assets (the user) is notoriously susceptible to social tactics such as deception, manipulation, and intimidation, and savvy threat agents know how to use this to their advantage" [71]. This is in line with the common saying in IT security, that the human being is the weakest link of the security chain [50, 64]. Social engineering bypasses every digital and physical security feature [64]. In this thesis, an attempt will be made to shed more light on this relation between humans as individuals and the success or failure of social engineering attacks.

This paragraph defines the scope of this work. Social attacks via persuasion and influence are considered, while direct threats including threats of violence, coercion, blackmailing, bullying, or bribery are not considered. Another differentiation refers to willingly or unwillingly, and intentional or unintentional acts from the targeted employee. The employee is not forced to comply through direct threats but decides to comply with the request of the attacker. Thus, whatever he or she does is done willingly. Both intentional and unintentional violations of security policies and disclosure of private information are within the scope of this work: At the time of the attack, the employee could either not think about violating policies intentionally but simply comply with a request, or he could intentionally decide to comply with the request, although he knows that this violates policies. Both cases are within the scope of social engineering attacks. In the first case, the targeted employees often question their actions afterwards, if they realize that they have done something that violated security policies or common sense. Nevertheless, it could also happen that the employee does not even realize that he has been the target of an attack or that he has disclosed private information. A possible distinction could be made regarding to competitive intelligence. Although the scope of techniques used for industrial espionage is much broader, it can and supposedly does also include social engineering attacks for information and financial gain. Thus, industrial espionage cannot be clearly confined from social engineering, but it can be seen on a meta-level. Market research can contrariwise be differentiated clearly from social engineering, since it only uses legal and publicly available means to information gathering about market and competitors. In this work, it will not be discussed how exactly detection of and defence against social engineering attacks work. Rather, some aspects, which can improve both, will be analysed.

The methodology used in conducting this investigation is a theoretical, positive research approach [54]. Extensive literature reviews have been conducted on social engineering, psychology of persuasion and the influence of personality traits on the success of social engineering attacks. Based on an analysis and discussion of these reviews, proposals for a new framework and a prevention approach have been developed.

In Chapter 2, the current state of research pertaining to social engineering is presented. In Chapter 3, Robert Cialdini's principles of influence and their relation to social engineering are discussed. While these principles are assumed to apply universally to every human being, not every person seems to be equally vulnerable to social engineering attacks, which suggests that individual differences in targets might mitigate the effects of persuasion attempts. In Chapter 4, the concept of personality, which is one such individual factor, and the leading theoretical model are presented. A framework is developed for how personality affects susceptibility to social engineering attacks in Chapter 5, supported by research results and supplemented with research proposals for hitherto untried relations. Knowing about the relationship between personality and social engineering susceptibility enables us to think about more specific prevention possibilities. These are discussed in Chapter 6. At the end, a conclusion and an outlook are provided in Chapter 7.

2 Social Engineering

2.1 Social Engineering Techniques

Generally, social engineering techniques can be differentiated into human-based and computer-based attacks. Human-based attacks require direct interaction (face to face or via telephone), while computer-based attacks, as the name suggests, usually take place in the digital world [55]. The second dimension inherent in this differentiation is direct versus indirect. Although human-based attacks per definition require direct interaction, computer-based attacks can be either direct as in chats or messengers, or indirect as in emails or websites. The direct computer-based attacks usually involve techniques subsumed under human-based attacks, such as impersonation or pretexting. Looking at the phases of a social engineering attack, one can differentiate pre-attack techniques for information gathering and the actual attack techniques. The pre-attack information gathering phase involves only indirect actions. Therefore, it does not include human-based attacks, since those are direct. It can, however, contain indirect computer-based techniques. The actual attack phase includes the classical social engineering techniques both human-based and computer-based. Table 2.1 visualizes the classification described above, and the following listing shows social engineering techniques that have been discussed in previous research (see, for example, [37, 69]), with the exception of direct computer-based techniques. The latter have not been discussed in the literature but are included for completeness regarding to the classification above.

	Human-Based	**Computer-Based**
Pre-Attack Information Gathering	-	Indirect
Classical Social Engineering Attacks	Direct	Direct & Indirect

Table 2.1: Dimensions and Phases of Social Engineering Attacks

Pre-Attack Information Gathering

- Techniques without contacting a target (Indirect)
  - Dumpster Diving (searching for information in the garbage of the organization)
  - Shoulder Surfing (observing people when they use confidential information such as login data)
  - Eavesdropping
  - other Observation Techniques (in person and technologically)
  - Stealing

- Computer-based Attacks (Indirect)
  - Social Network Information Gaining
  - Website Surfing (open source information)
  - Guessing Passwords

Classical Social Engineering Attacks

- Human-based Attacks (Direct)
  - Impersonation and Pretexting (e.g. posing as an insider or legitimate visitor)
    - Appealing to Emotions
    - Being Likeable
    - Establishing Rapport
    - Flirting
    - Lying
    - Manipulation & Deception
    - Name dropping
    - NLP (Neuro-Linguistic Programming, a technique using gesture, facial expression, tone of voice, language, and keywords to gain compliance)
    - Reciprocation
    - Social Validation
    - Using Authority
    - Using Fake ID
    - Using Jargon or Lingo
  - Tailgating / Piggybacking (“Please hold the door!”)
- Computer-based Attacks (Indirect)
  - Phishing
  - Placing USB devices with malicious code inside (or in the vicinity of) the company
  - Malicious E-Mail Attachments
  - Using Key-Loggers
- Computer-based Attacks (Direct; uses human-based techniques)
  - Chat-based
  - Messenger-based
  - Video-based

Concerning information gaining via social networks, Huber et al. introduce an attack called automated social engineering, where bots are used to collect information freely available in a social network, and to directly contact people via social networks to elicit information [36]. This enables attackers to gain even more information with less personal effort, and increases the probability of social engineering being used even further.

2.2 Defining Social Engineering

The first thing noticed in a research about social engineering is that most publications cite non-scientific and popular scientific sources rather than scientific publications for defining the subject, which is probably due to the fact that the topic was first recognized by the public and economy after some famous (or rather infamous) individuals like Kevin Mitnick or Frank W. Abagnale, Jr. had publicized their past activities as social engineers [1, 41, 48]. Only after the public and the economy gave the topic a wide range of attention, IT scientists started to engage in research about social engineering as a distinct subject. Although nowadays there are quite a few scientific publications about social engineering, the most comprehensive compilations about this topic still seem to stem from websites such as `http://www.social-engineer.org` or `http://www.csoonline.com`. These websites take a very practice-oriented view, and offer many examples and proposals for prevention.

Hadnagy, one of the creators of `http://www.social-engineer.org`, defines social engineering as “the act of manipulating a person to take an action that may or may not be in the target’s best interest. This may include obtaining information, gaining access, or getting the target to take certain action” [31]. He has proposed a social engineering framework that includes a collection of skills that “when put together make up the action, the skill, and the science” [31] he calls social engineering. The framework defines consecutive phases of a social engineering attack: information gathering, elicitation, and pretexting. Furthermore, he specifies psychological principles and different tactics of influencing others. Lastly, he presents physical, computer-based, and phone tools for social engineering. Hadnagy’s framework is suggestive of a patchwork that combines everything that has been associated with social engineering over the years. Additionally, the phases, which are presented, appear to be hard to differentiate: both elicitation and pretexting can have the primary goal of information gathering. Hadnagy takes a very broad approach in his definition of social engineering, as he includes legitimate persuasion episodes like psychotherapy or educational situations, and even children trying to get their parents to do something. Hadnagy justifies this broad approach with the argument that although the settings and intentions are diverse, malicious attackers and legitimate users of social engineering, like doctors, psychologists or teachers, are using the same approach. However, with respect to social engineering as a research field of IT security, the focus is on a malicious social engineer trying to gain confidential information for breaking into a system. Thus, we need a more specific approach for a usable definition.

Therefore, a literature review has been conducted to compare definitions of social engineering. While there seems to be no universalized definition commonly used, there are certain aspects that

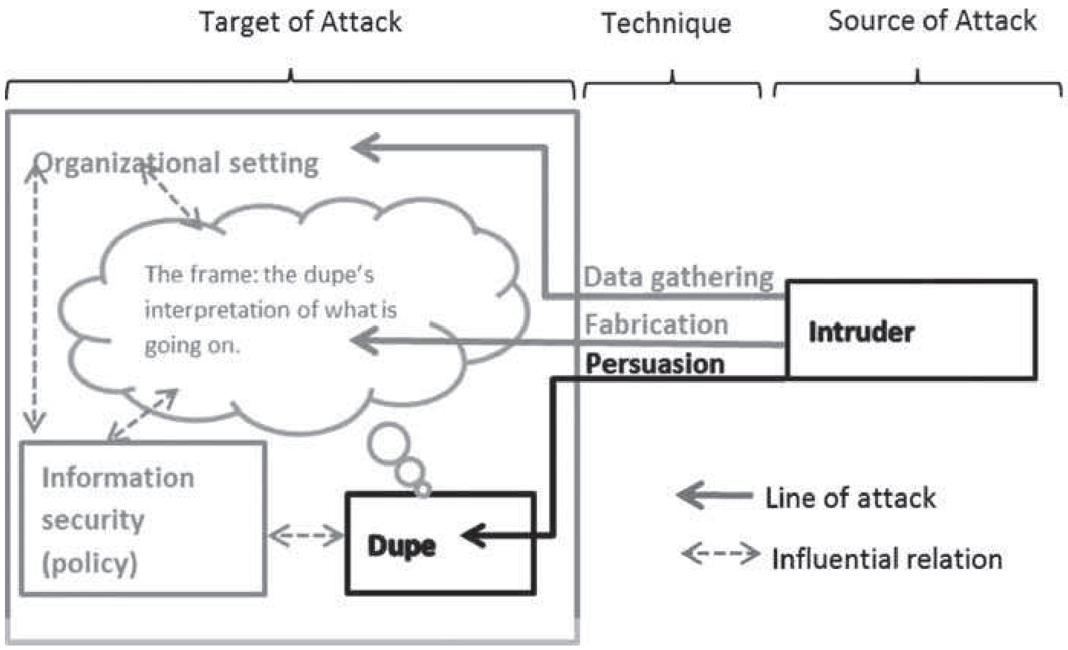

Figure 2.1: Tetri et al.'s Social Engineering Framework [69]

most publications refer to (see, for example, [2, 37, 49, 50, 62, 69]). Generally, two actors are defined: an attacker, the social engineer, and a victim, mostly referred to as the target. The attacker uses some kind of techniques to make the target perform actions or divulge confidential information (for example, a user name and password) that he or she would not have performed or divulged normally. Implied in this approach is some sort of trickery or persuasion to make the target comply. Many authors explicitly include an exploitation of human traits like trust and emotions by the attacker. The goal of a social engineering attack is mostly stated as getting access to confidential information or whole IT systems, usually for sabotage, financial gain or identity fraud. The reason for using social engineering in the first place is that most attackers perceive social engineering as easier and more promising compared to searching for technological weaknesses and hacking into systems, due to the fact that in the past, the focus of IT security has been on technological defence against perpetrators, leading to better perimeter defences, which increase the difficulty of breaking into a system technologically. Social engineering, however, requires only a minimum of technical expertise. Moreover, it uses an attack vector that cannot simply be patched when a vulnerability is detected – the user.

Tetri et al. declare that most of the social engineering literature focuses on individual techniques that were used in incidents. They state that this focus leads to a "scattered, anecdotal, and vague notion" [69] of social engineering, and a lack of analytical concepts. The consideration of single incidents without an extraction of the basic principles and mechanisms has lead Tetri et al. to the conclusion that “the victim’s psychological traits are overemphasized, although this kind of explanation can cover only a small portion of social engineering cases” [69]. Therefore, they have developed a comprehensive framework of social engineering, that does not only focus on the attacker-target-relation but also encompasses situational and organizational conditions. Through the analysis of social engineering techniques, they extrapolated three dimensions of social engineering: persuasion (“getting a person to comply with an inappropriate request” [69]), fabrication (“providing misleading cues to the target in order to affect its interpretation of what is going on in the situation” [69]), and data gathering (getting information for further intrusion, not necessarily based on direct interaction). Tetri et al. argue that through these dimensions, social engineering can be described in all its aspects. Every social engineering technique can be characterized by its manifestation of these three dimensions, whereat usually not a single dimension but a mix of all three dimensions applies. In Figure 2.1, Tetri et al.’s framework of social engineering is shown. The elements that they argue to be overemphasized in the previous literature are displayed in black, while the grey elements are at most implicitly present in the previous literature but are nevertheless stated to be essential in analysing social engineering. Tetri et al. argue for a multidimensional approach for analysing social engineering. Thus, in their framework, the relation between attacker and target has been expanded to also include organizational settings and the information security policy of the organization, both of which influence the target’s interpretation of the situation. This approach seems sensible since it broadens the scope of analysis, and gives more starting points for prevention.

Comprising the points discussed above, I define social engineering adopted for this work as followed:

Definition 1. *Social engineering is an attack where a human is illegitimately manipulated into performing actions or divulging information that he or she would not have done without the manipulation. To achieve this manipulation, the attacker uses a set of techniques to be more convincing, including persuasion as well as impersonation, information gathered beforehand, or cues, which leads the targeted employee to misinterpret the situation according to the attacker’s intent. The interpretation of and reaction to the attack by the target is also defined to belong to social engineering, since the dynamic interaction between attacker and target is a central part of social engineering. The target’s interpretation of the attack, including the realization that an attack is happening at all, is influenced by situational and organizational settings as well as individual experience. The goal of the social engineer is getting access to confidential information or systems for sabotage, financial gain, or identity fraud.*

According to this definition, an essential part that determines the success of a social engineering attack is the target’s ability to detect the attack and resist the manipulation. The following chapters discuss important aspects that influence these abilities.

3 Cialdini’s Psychology of Persuasion and Related Theories

Most people think that they are not likely to be the target of a social engineering attack, and moreover that they can detect social engineering attacks and persuasion attempts, respectively (see e.g. [60]). However, this is provably not the case, otherwise, there would not be that many reports of successful social engineering attacks. Where do these misjudgements come from? Schneier [65] argues that most wrong appraisals of risk are evolutionary justified. Human risk perception has evolved over thousands of years, and has worked well until the pre-industrialized age. However, technological progress has changed our way of living so fast that the slower evolutionary process has not had time to adjust. “Like a squirrel whose predator-evasion techniques fail when confronted with a car,[...] our innate capabilities to deal with risk can fail when confronted with such things as modern human society, technology, and the media. And, even worse, they can be made to fail by others — politicians, marketers, and so on — who exploit our natural failures for their gain” [65]. Table 3.1 shows the most common risk misjudgements. Some of these entries can explain why people underestimate their risk to be a target of social engineering, e.g. it is anonymous, and not much discussed in the public. However, this approach cannot justify all misjudgements of risk, especially not those pertaining to social life, since social life exists as long as the human race. The social engineer usually tries to be perceived as a trustworthy person, for example, a member of a peer group, when posing as an insider. Thus, he circumvents most risk assessments that relate to technological progress, which implies that the evolutionary approach to risk misjudgements is not adequate for this specific scenario.

Another point discussed by Schneier as a reason for wrong appraisals and decisions are heuristics. Heuristics are short-cuts, stereotypes, rules of thumb, and biases that we use to reduce cognitive load [65]. These heuristics are evolutionary beneficial, just like risk perception. However, just like risk perception, some of these heuristics are just not suited to modern life, and can be exploited by others. Admittedly, in the complexity and with the information overflow of modern life, it is just not possible to live without heuristics. People cannot fully analyse every decision. Cialdini [10] discusses how a malicious person can take advantage of the heuristics and biases we have. For this, he has deduced six principles of influence, which will be presented in the next section. Later in this chapter, it will be shown that his work, which is mainly about sales and marketing, can be adopted to social engineering and the biases and heuristics we have about risk and security.

Cialdini uses the metaphor of a „Click-Whirr“-behaviour [10], depicting an automated behaviour usually triggered by a single feature of the relevant information in the situation (“Click”), the heuristic.

People exaggerate risks that are:	People downplay risks that are:
Spectacular	Pedestrian
Rare	Common
Personified	Anonymous
Beyond their control, or externally imposed	More under their control, or taken willingly
Talked about	Not discussed
Intentional or man-made	Natural
Immediate	Long-term or diffuse
Suddenly evolving	Slowly over time
Affecting them personally	Affecting others
New and unfamiliar	Familiar
Uncertain	Well understood
Directed against their children	Directed towards themselves
Morally offensive	Morally desirable
Entirely without redeeming features	Associated with some ancillary benefit
Not like their current situation	Like their current situation

Table 3.1: Conventional Wisdom About People and Risk Perception, taken from Schneier [65].

It helps to decide on correct action ("Whirr") without having to completely analyse the whole situation in every detail. This is efficient and economic because it saves energy, time, and mental capacity. The disadvantage is clearly the possibility that by reacting only to one piece of the available information, errors can happen, especially if someone is exploiting this trigger information intentionally.

The theory that is at the base of this automated short-cut responding is the Dual Process Model of Persuasion, also known as the Elaboration Likelihood Model [30]. It states that there are two different ways in which humans process information: centrally and peripherally. The central route, or systematic processing, is taken when the individual focuses on the content of the message. Decisions are made based on qualitative factors of the arguments. When the peripheral route, or heuristic processing, is taken, the individual uses heuristics to decide on his or her attitude on the topic. There is evidence for some factors making peripheral processing more probable, including strong affect, lack of motivation, lack of personal relevance of the topic, lack of knowledge about the topic, lack of cognitive ability to process the message, lack of time, cognitive comfort due to trust, and communication modes where the influence agent is salient [6, 10, 30, 60, 77]. However, even if a topic is important to us, we are not always able to make decisions based on a full analysis due to the pace and complexity of modern life [10]. As a matter of fact, it is quite probable that most decisions are made without full consideration or even full knowledge of all facts.

Social influence refers to "the change in one's attitudes, behaviour, or beliefs due to external pressure that is real or imagined" [30]. There are two specific types of influence that are of interest in this work: compliance and persuasion. Persuasion focuses on the change in attitude, knowledge or belief as a result of a received message. Compliance focuses on change in behaviour that results from a direct request. "The request may be explicit [...] or it may be implicit, [...] but in all cases, the target recognizes that he or she is being urged to respond in a desired way" [11]. Compliance and persuasion

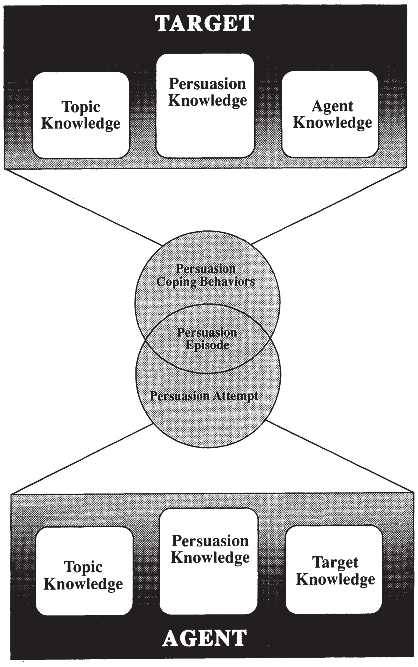

Figure 3.1: The Persuasion Knowledge Model by Friestadt and Wright [26]

can both be understood in terms of a human tendency for automatic shortcut responding. The set of trigger features (heuristics) for compliance tells us when compliance is likely correct and beneficial, which is what is usually abused by a social engineer. While persuasion in this work is regarded as interchangeable with influence, and represents the action of the social engineer, its effect can be both persuasion and compliance with respect to the definition above.

Friestad and Wright have developed a persuasion knowledge model (see figure 3.1) that includes the points of view of both parties in a persuasion episode [26]. Persuasion knowledge is defined as the personal knowledge of an individual about persuasion attempts, which helps to identify how, when, and why someone tries to influence him or her. Friestadt and Wright assume that people's goal in a persuasion episode is to maintain control over the outcome, and achieve their salient goals. Therefore, the roles of target and persuasion agent are fluent, and can alternate multiple times during an episode. The use of persuasion knowledge in this situation is not necessarily typical. Persuasion coping behaviour of targets encompasses their cognitive and physical actions during the episode as well as their thinking about an agent's persuasion behaviour before, between, and after episodes. Friestad and Wright propose three relevant knowledge structures that interact to determine the outcome of a persuasion attempt: persuasion knowledge, agent knowledge (what the target believes about the traits, competencies, and goals of the persuasion agent), and topic knowledge (beliefs about the topic of the message). The target's persuasion-coping knowledge is proposed to enable the target to recognize, analyse, interpret, evaluate, and remember persuasion attempts as well as to select and execute coping tactics believed to be effective and appropriate [26]. It is assumed that the development of persuasion knowledge depends on the maturation of some basic cognitive skills and on people's accumulated experience with what occurs in social encounters as well as social discourse about persuasion, advertising, and psychological events. Thus, a practical application of this model in prevention could be based on the building of persuasion knowledge through awareness trainings and personal experience with social engineering attacks. More on this will be discussed in chapter 6.

3.1 The Six Principles of Influence

The principles of influence are assumed to generally apply to every human being. They are all subsumed under Cialdini's "Click-Whirr"-automated-behaviour. Cialdini extracted them by experimental studies and by field studies in the world of influence practitioners, predominantly in marketing and sales. The crux with these principles is that even if one knows about them, it is not easy to identify illegitimate use. As the behaviour triggered by these principles is an important source for social coherence, and is deemed positive and socially desired behaviour, it is not advisable to reject it altogether. Rather, it is needed to be able to identify illegitimate usage of fabricated triggers to specifically cope with it. For the following subsections, where no other citation is made, the content refers to Cialdini's book "Influence: science and practice" [10].

3.1.1 Reciprocation

Reciprocation is a very strong social norm that obliges us to repay others for what we have received from them [11]. It is so strong because our society is based on it. Without it, continuing relationships and exchanging goods and services would never have been developed in human history. Reciprocation helps building trust with others, and refers to our need for equity in relationships. Every child is trained to adhere to this norm or suffer severe social disapproval, which explains why reciprocation is such a powerful principle. The pressure of obligation can be so high that to get rid of it, the target will pay back a greater favour than he or she received before. The reciprocation principle works with concessions as well. Even uninvited first favours invoke an obligation. The exception is a favour that is clearly identified as a leverage. This is perceived as an illegitimate tactic, which usually backfires, and stirs reactance. The best defence against illegitimate use of reciprocation is to accept initial favours and concessions but be ready to re-evaluate them as tricks whenever they prove to be such, so no feeling of obligation is aroused.

This principle is used in social engineering by offering free gifts or favours in advance of the attack, thus increasing the probability of compliance due to a feeling of obligation. [15]

3.1.2 Commitment and Consistency

Will Durant abstracts an idea of Aristotle as: "We are what we repeatedly do" [19]. This quotation is a very good description of the principle of commitment and consistency. They are essentially two sides of the same coin: Commitment is an act of stating what one person thinks he is and does, while consistency makes that same person behave consistently according to his or her commitments, beliefs, and self-ascribed traits [11]. "This consistency with prior commitments is a highly successful influence principle because it alters one's self-perception. We often look to our own behavior to understand who we are. However, the outcome of our actions based on self-perception information varies based on the level of internal consistency we desire and the way a request is presented" [30]. However, there seem to be some conditions on commitments and actions to alter our self-perception: they need to be active, public, effortful, and freely chosen. Furthermore, the individual level of internal consistency we aspire, and the way the request is presented also have some influence.

The theoretical basis for this principle is Festinger's and Carlsmith's cognitive dissonance theory [22, 77]. It proposes that people are motivated to maintain congruence between their attitudes and behaviours to minimize the experienced cognitive dissonance.

Cialdini et al. [13] constructed a scale for measuring preference for consistency. They were able to show that preference for consistency exists as a measurable personality trait, and that the 'Preference for Consistency Scale' is a valid instrument for assessing it. The study revealed a surprising insight: more than half of the participants held no particular natural preference for consistency. In another study, they found that preference for consistency, and subsequently, the likelihood to comply with techniques using this principle, increases with the years and is strongest in people older than 50.

Cialdini suggests that the best defence against this principle is to listen for signals from our body that reacts with unease whenever we are being pushed by this principle's pressures. But as Schneier [65] has shown, our feelings and risk perceptions do not work that well in modern life.

One of the most widely used techniques using the commitment-and-consistency-principle is the foot-in-the-door technique: the target is made to comply with a seemingly small request. After the target committed him- or herself to comply, the real request is presented. The likelihood to comply with this, usually essentially bigger, request is much higher after an initial commitment to a related request. In social engineering, this could be used in a variety of ways. The small request could pertain to being helpful in general, so the likelihood of helping with the real request is heightened. This principle can also be used to attribute certain characteristics to the target, e.g. telling him at the beginning how widely they are known for their cooperativeness and helpfulness, which can connect the target's identity to the subsequent request by giving them a reputation to uphold.

3.1.3 Social Proof

Social Proof is all about doing or believing what the other people around a person are doing or believing, especially in situations that are ambiguous, and if the others are similar to that person. It also implies that we trust people who are like us, for example our friends, and their decisions and actions.

Ryan [59] states that the position his false identity Robin Sage claimed to occupy led people to assume that she had passed background checks of the government, thus implying some level of trust that could be given to her. This could of course be subsumed as belief to authority, but it can also be seen as evidence that other relevant people from the field of security trust her, so, by social proof, oneself can also trust her. Even (self-proclaimed) experts in the field of Cyber-Security connected to her without any check of her background and trustworthiness. The more people from the field of IT-Security connected to her, the greater became the impact of Social Proof that the fake accounts could use for connecting to other people. This was enhanced by Robin befriending some very respected specialists in the security sector. Ryan also states that social networks "encourage people to grow their network by establishing trust based on mutual friends. [...] Cyber predators are aware of this ’mutual friends’ oversight and exploit it readily” [59].

Cialdini [14] conducted a study on both social proof and commitment-consistency. Both principles were influential across cultures, but the commitment-consistency principle had a greater impact on Americans, whereas the social proof principle had greater impact on Poles. This affect was attributed to differences in individualistic-collectivist orientations of the individuals and their culture respectively.

The suggestion to resist the usage of this principle is to challenge evidence of social proof regarding its validity, and to not form decisions based solely on social proof arguments.

In social engineering, this principle could be used by mentioning friends or colleagues of the target that have supposedly told the attacker how helpful or knowledgeable the target is, thus using both the principles of social proof and commitment-consistency at once. Alternatively, the attacker could state how many colleagues have already complied with his request, for example, filling out a survey including user name and password. Of course, this only works if this statement makes sense in the scenario used by the attacker.

3.1.4 Liking

“If you make it plain you like people, it’s hard for them to resist liking you back” [9]. This quote from Bujold describes the liking principle perfectly. We prefer to comply with requests from people we know and like. The reason for this is the fundamental human motive to create and maintain social relationships with others. It takes as little as perceived similarity to enhance compliance, even if those similarities are as superficial as shared names or birthdays, since similarities are cues for a potential friend. This effect may be stronger among women than men, since females tend to be more relationship-oriented than males [11]. Another factor that enhances liking and consequently compliance to requests as well as changes in others’ attitudes is physical attractiveness.

Ryan emphasizes “the role that sex and appearance plays in trust and people’s eagerness to connect with someone” [59]. He deliberately chose “a young, attractive and edgy female” for his false identity, arguing that in the predominantly male sector of IT security, being a woman would heighten the interests, and make people more comfortable. This assumption was confirmed by many comments on Robin Sage’s social networking sites and personal messages. Another aspect of liking experienced by Ryan was the common ground of educational ties. Many people connected to the fake account on the basis of having been at the same school or college without ever checking if Robin even attended them.

Workman comments on a finding that “people usually ’trust those they like’, and conversely, they usually ’like those they trust’” [77].

Social engineers often use humour to create liking towards them. They also often play on employees’ desire to be helpful [15]. This can be seen as liking in reverse: we all want to be liked, and perceived as helpful.

3.1.5 Authority

Authority is probably the most plausible principle, since most people have made experience with complying to authorities during their lives. The most famous scientific experiments about obedience to authority are the Milgram-Experiments [47]. They show that use of authority may even let us act against our beliefs and ethics. This usually works for symbols of authority as well, e.g. uniforms, badges, ID-cards, and titles. Authority is especially easy to forge as the symbols for authority can be fabricated rather easily, and for telephone conversations, authority can simply be claimed, and is hard to challenge. Moreover, people tend to trust people whom they perceive as credible regarding special expertise or authority [77]. Another case in point for the success of this principle is the real story about the "Hauptmann von Köpenick", where a former convict, wearing the right uniform, posed as a captain of the Prussian army, commanded a troop of soldiers, imprisoned the major of Berlin Köpenick and stole the city treasury [73].

There are two different types of authority, one based on expertise and one based on the relative hierarchical position in an organization or society. They correspond to soft and harsh influence techniques, the former being based on factors within the influence agent, and the latter being based externally on existing social structure [11]. In social engineering, both types of authority are used. The expert authority is often used when the attacker impersonates someone from the IT security department, claiming that he needs ID and password of the target due to some problem with the target's computer. The hierarchical authority is often assumed to create an ambience of importance and stress, for example, when the social engineer calls a help desk, and claims to be some manager who has lost his password but very desperately needs access to his e-mails.

3.1.6 Scarcity

We assign more value to opportunities that are less available. This is due to a short-cut from availability to quality. Moreover, if something becomes less available, we lose freedoms. Reactance Theory suggests that we respond to this loss of freedoms by wanting to have what has become increasingly rare more than before [7]. Another interesting finding is that limited information, in this case information that is not generally accessible, is more persuasive.

Scarcity is probably the principle that is hardest to transfer to social engineering. In sales, it is very easy to say that there is only a limited quantity available. However, it can be used in computer-based social engineering to heighten the pressure to download programs or click on links. For example, if the attacker knows what kind of software would be very useful for the target, he could lure the target with an e-mail or advertise a free trial on a website that is only available for a few hours or for a limited number of users (First-come, first-served). This can greatly enhance the probability of download of the (malicious) software that enables the attacker to compromise the computer and probably the whole system.

3.2 Persuasion and Compliance on the Internet

Since persuasion and compliance are an interactive phenomenon, it is reasonable to assume that on the internet, a medium without direct personal contact, some aspects and results could be different. Guadagno and Cialdini have written a review article about this topic [30]. For the following section, where no other citation is made, the content refers to their article.

There are some aspects of online interactions that should be considered, because they might influence how persuasion and compliance work online. The internet enables people to stay anonymous almost as much as they like. Since most online interactions are text-based, physical appearance is less important than in direct face-to-face conversations. This extends to the absence of non-verbal communication in general. Nowadays, it is technically feasible to provide real-time video conversations, but since this is not standard in today's business communication, this case will not be considered here. In this context, the person trying to get the target to comply is called the influence agent. In a social engineering context, this would be the attacker.

Since most physical and communicator cues such as liking, status, and expertise are less salient in an online interaction, their impact on persuasion and compliance should be less important compared to face-to-face interactions. Then again, social category cues can still be available. If the target is informed that the influence agent is of a certain profession or social or organizational status, this might influence his response. Photographs, for example in interactions on social networks, were shown to increase compliance for both men and women, although the effects were stronger for men. They provide social status cues. Both men and women were more compliant when the influence agent was female.

Generally, Guadagno and Cialdini have found evidence that messages received through the medium computer are more likely to be centrally processed, especially if the persuasive attempt is non-interactive as for example in e-mails. A gender effect was reported: females show less agreement with a message if it is presented via e-mail, compared to a face-to-face communication. This held regardless of the strength of the arguments. Males contrarily showed no communication mode difference. The explanation offered for this effect was a relationship-orientation of females, whereas males were oriented towards the task and their own independence.

Only two of the six principles of influence have been explicitly examined empirically regarding the internet: authority and commitment-consistency. As expected, status and expertise were less salient in computer-mediated decision groups. Authority was shown to increase compliance in online groups if used as a decision heuristic. However, in an interactive discussion its influence was much smaller. Regarding commitment and consistency, the foot-in-the-door technique was shown to be effective in several computer-mediated contexts as a function of desire for internal consistency.

While many studies examining phishing found demographics to be a mediator for phishing susceptibility, Mohebzada et al. contradicted these results in their large scale phishing experiments [49]. Workman emphasizes the nature of online-trust and its importance for conducting online business transactions, but also warns that it can increase vulnerability for social engineering attacks [77].

In aggregating the results, Guadagno and Cialdini state "as a general rule, influence appeals that are mediated by self-focus should operate similarly in online contexts [...], but others that rely on an interpersonal interaction may function differently in an online interaction" [30].

3.3 Mapping Social Engineering to the Principles of Influence

Tetri et al. question the existing research on social engineering on the topic of uncritically assuming validation of psychological theories in another field as they have been developed for [69]. Specifically, they challenge the assumption that Cialdini's principles of influence, which have been written from a perspective of sales and marketing, can be adopted to IT security. However, as was shown above, social engineering attacks using each of the principles are possible. Moreover, Scheeres argues in his thesis that these principles can be used in the context of social engineering. In his reasoning, he compares Gragg's psychological triggers, which are explicitly referred to as being targeted by social engineering attacks [29], with Cialdini's principles of influence to evaluate whether social engineering can be equated to (illegitimate) persuasion [62]. In Table 3.2, the result of this comparison is displayed.

Cialdini's Principles of Persuasion	Gragg's Psychological Triggers of Social Engineering
Scarcity	Strong Affect
	Overloading
Reciprocation	Reciprocation
Liking and Similarity	Deceptive Relationships
	Diffusion of Responsibility
Authority	Authority
Commitment and Consistency	Integrity and Consistency
Social Proof	

Table 3.2: Comparison of Cialdini's Principles and Gragg's triggers. Taken from Scheeres [62].

Gragg states that the arousal of strong emotions are used in many social engineering attacks to disable the target's central route processing. He calls this trigger strong affect. Scheeres argues that nearly all applications of the scarcity principle fit into this trigger, but that the trigger can be more broadly applied than the principle [62]. Gragg's second trigger, overloading, also describes tactics to disable central route processing by overwhelming the target with arguments in quick succession. This trigger cannot be matched to any principle of influence. Reciprocation is the same in both approaches. The trigger of deceptive relationships describes that in social engineering, the attacker often makes the target believe that a relationship between them exists, or even builds such a relationship before stating his request. This trigger and the principle of liking are based on attributes that the target desires, thus, they can be seen as interchangeable. When using the diffusion of responsibility trigger, a social engineer makes the target believe that he or she is not solely responsible for his or her actions. This can be done, for example, by hinting that someone higher up in the hierarchy has already made the decision, and that the target simply has to execute this decision. This trigger could be mapped to authority, but since authority is defined as a trigger on its own that perfectly matches the corresponding principle, this does not make sense. Instead, this trigger focuses on making the target believe that an action that clearly violates security policies will have some greater benefit for the organization. In this regard, there does not seem to be an equivalent principle of persuasion. The trigger of integrity and consistency matches the principle of commitment and consistency. Lastly, the principle of social proof does not have a correspondent trigger in Gragg's taxonomy. Scheeres concludes that "while illegitimate persuasion and social engineering cannot be exactly equated, the principles of persuasion are very similar to the psychological triggers of social engineering" [62]. Due to this statement and to the previously shown exemplary social engineering attack for every principle of influence, it is concluded that Cialdini's psychology of persuasion can be used as a valid theoretical basis for social engineering.

4 Personality Traits

In psychology, personality is defined as a person's relatively stable feelings, thoughts, and behavioural patterns. These are predominantly determined by inheritance, social and environmental influence, and experience, and are therefore unique for every individual [3].

The most common classification approach in personality psychology is to try to extrapolate as few statistically independent, usually bipolar dimensions as possible. The first classifications of this kind appeared in the 1950s [3]. Every dimension can be measured by scales that correspond to variables in which people differ. These dimensions are labelled as personality traits, and are defined as relatively stable dispositions that manifest across different situations and a certain space of time. Of the classification approaches, the Big 5 Theory, also known as the five-factor model [43] has established itself as most widely used and extensively researched theory, although it does not enjoy universal agreement.

4.1 The Big 5 Theory of Personality

As the name suggests, the Big 5 Theory of Personality consists of five broad, empirically derived personality dimensions or traits, which are sufficiently independent of each other, and are believed to capture most of the individual differences in personality. These dimensions are conscientiousness, extraversion, agreeableness, openness, and neuroticism. Each of these factors can be split in several sub-traits. These traits have been used to predict behaviour across a variety of situations and areas of research with high validity.

High values in the five traits are described as follows [43]: Conscientiousness encompasses a focus on competence, self-discipline, self-control, persistence, and dutifulness as well as following standards and rules. Extraversion comprises positive emotions, sociability, dominance, ambition, and excitement seeking. Agreeableness includes compassion, cooperation, belief in the goodness of mankind, trustfulness, helpfulness, compliance, and straightforwardness. Openness to experience is defined as a preference for creativity, flexibility, fantasy as well as an appreciation of new experiences and different ideas and beliefs. Neuroticism is the tendency to experience negative emotions, anxiety, pessimism, impulsiveness, vulnerability to stress, and personal insecurity. Conversely, low values in these traits represent the opposite of these attributes.

Shropshire et al. argue that the five-factor model is best suited to be used in a context of IT security, since it is a generalizable taxonomy that permits use across many different research disciplines. Moreover, the behavioural patterns that are associated with its factors are more extensively researched than any other more specific personality factors [67].

Hirsh et al. discuss evidence that people with high values in specific traits have different motivational systems [33]. Thus, people with high values in extraversion are motivated by rewards and social attention. High values in agreeableness correspond with communal goals and interpersonal harmony. Conscientious individuals are motivated by achievement, order, and efficiency. People with high values in neuroticism are sensitive to threats and uncertainty, while openness corresponds with creativity, innovation, and intellectual stimulation [33]. These motivations could be capitalized by social engineers as well as in preventive measures.

4.2 The Business-Focussed Inventory of Personality

One aspect of most questionnaires that measure personality traits is that they enable the creation of a comprehensive personality profile, including implications of personality disorders. In Germany, the use of such questionnaires in the context of employee selection is controversial due to the general right of personality (article 1 and 2 of the German Constitution). The right of the individual to protect and develop his personality is opposed to an employer's interest to find the best employee including his personality. Business-related appraisals of attitude are unobjectionable [32].

One such business-related instrument is the Business-Focussed Inventory of Personality (BIP). It consists of four broad scales: occupational orientation, occupational behaviour, social competencies, and psychological constitution. Each of these scales consists of several sub-scales. Occupational orientation covers the work-specific motivation. It involves planning and shaping the career path as well as values in a job. Sub-scales are achievement motivation, power motivation, and leadership motivation. Occupational behaviour measures the typical approach to work and focuses on the sub-scales conscientiousness, flexibility, and action orientation. The scale of social competencies displays the style of interaction with other people, including the sub-scales social sensitivity, openness to contact, sociability, team orientation, and assertiveness. Lastly, psychological constitution describes how one reacts to demands made by work with respect to resilience and the experience of emotional pressure. The sub-scales are emotional stability, working under pressure, and self-confidence. There is a certain overlap with this inventory's sub-scales and the Big 5 Theory.

The inventory can be seen as a both reliable and valid instrument to measure business-oriented personality traits. The scales have medium to high correlation to other scales for personality [35], including the personality test NEO-FFI, which explicitly uses the Big 5 model [44]. This implies construct validity, meaning that the inventory really measures business-related personality traits.

A bonus of the inventory is its standardization, which has been conducted with samples of the working population in Germany [35]. As a result, there exist normative descriptions of average personality characteristics in different occupational groups. These job profiles or individual values of employees' personality could help to develop custom measures for prevention.

5 The Social-Engineering-Personality-Framework

As was shown before, Cialdini's psychology of persuasion can be used as a valid theoretical foundation for social engineering. Thus, the framework, which is presented in this chapter, will propose relations between certain personality traits and a higher or lower vulnerability to social engineering mediated by the principles of influence. First, a simple model based on existing research is presented, together with a short literature review on this research. In the next section, a refined model will be proposed together with suggestions on how to provide evidence on the proposed relations. Both framework variants have been developed by the author of this work.

5.1 Basic SE-Personality-Framework and Research Evidence

Figure 5.1 shows the relations between personality traits and susceptibility to persuasion and social engineering, based on existing research. Several relations between personality traits and social engineering have been investigated in previous research. Most research explicitly used the Big 5 Theory, while others used narrower personality constructs out of which some have been shown to correlate to Big 5 traits. The results are rather broad: 3 out of 5 personality traits (Conscientiousness, Extraversion, and Openness) have been shown to both increase and decrease susceptibility to social engineering in different contexts and sub-traits. Agreeableness has been found to increase, and Neuroticism to decrease susceptibility to social engineering. The research pertaining to these rather diverse findings will be subsumed for each trait in the following.

Conscientiousness. Workman shows that continuance commitment, which is related to conscientiousness amongst other traits, increases social engineering vulnerability [77]. Opposed to this, Darwish et al. state in their survey of recent studies about phishing attacks and related backgrounds of victims that conscientious people who are more mature and show respect for standards and procedures have a lower rate of security risk [16]. Parrish et al. argue that this only applies to standards and procedures that are existent as well as communicated [51]. They also declare that security training should decrease social engineering susceptibility especially strong for conscientious individuals [51]. This is supported by research from Sagardo et al. where low levels of conscientiousness predicted deviant workplace behaviour such as breaking rules or generally behaving irresponsibly [61].

Extraversion. Darwish et al. state in their survey, that individuals with high values in extraversion are at a higher rate of security risk [16]. McBride et al. show that extraverted individuals are more likely to violate cyber-security policies [42]. Workman investigated the effect of different types of commitment to social engineering susceptibility. He ascertains that people with high affective com-

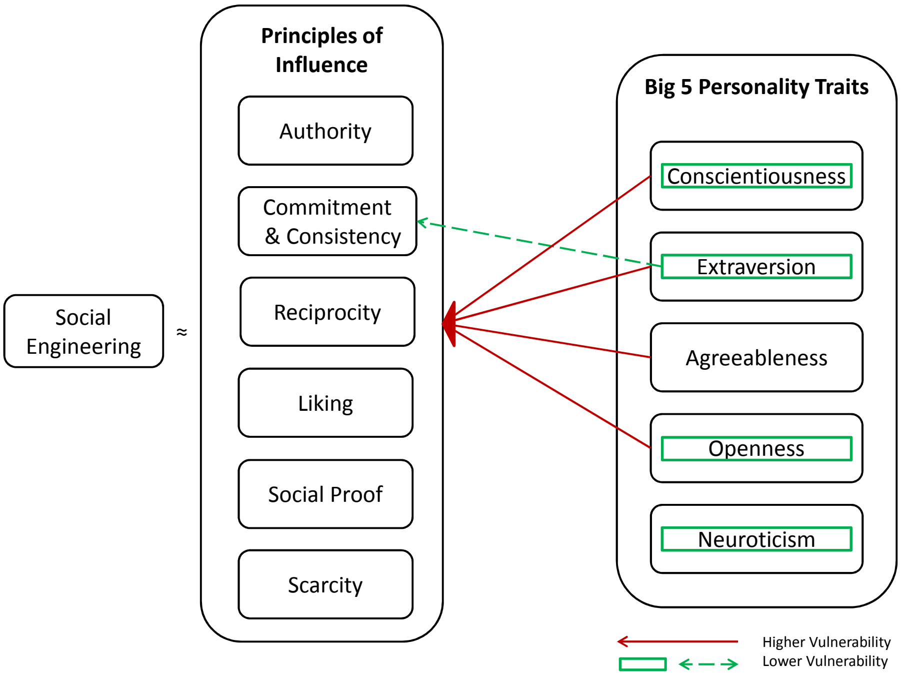

Figure 5.1: Simple SE-Personality-Framework based on existent research

mitment as well as high normative commitment were more likely to fall prey to social engineering attacks [77]. Both types of commitment have been shown to significantly relate to extraversion [20]. On the other side of the scale, Weirich and Sasse [75] report that employees who did not disclose their passwords, thus showing a low level of social engineering susceptibility, were regarded as unsociable and loners by their colleagues, implying low extraversion values. Controversially, Cialdini et al. show that people who are rated low on the preference-for-consistency-scale, thus being less vulnerable to commitment-and-consistency-techniques, show a greater extraversion than those high on the scale [13].

Agreeableness. Parrish et al. state that agreeableness is "possibly the personality trait that is most associated with" phishing [51], and in a greater scope social engineering. Darwish et al. report that individuals who are more agreeable are at a higher rate of security risk. They note that generally, younger people and women are known to have higher values of agreeableness [16], thus explaining some of the demographic differences found in phishing susceptibility. The relation between agreeableness and social engineering susceptibility is assumed to be mostly established by trust, a sub-trait of agreeableness. This was shown in studies by Weirich and Sasse as well as by Workman [75, 77]. In the latter study, high normative commitment – as written above – has been shown to increase social engineering vulnerability. It significantly relates to agreeableness just like to extraversion [20, 77]. Other sub-traits that have been found to be directly targeted by social engineers are altruism and compliance [51]. Sagardo et al. contradict these findings: they found that low levels of agreeableness predicted deviant workplace behaviour such as breaking rules [61]. This could hint at some interaction or constructional problem that should be examined in the future.

Openness. Junglas and Spitzmuller [38] reported that people with high openness values were less concerned about privacy problems associated with location based services. They argue that these people's tendency to seek new experiences influences their risk evaluation. This can be conveyed to social engineering in that open individuals underestimate their risk in becoming the target of such an attack, and subsequently do not develop adequate coping strategies. Controversially, McBride et al. found that more open individuals are less likely to violate cyber-security policies [42]. However, this effect is contradicted when personality is not evaluated as direct influence but as a moderating factor. In this case, open individuals have been found to be more likely to violate cyber-security policies [42].

Neuroticism. McBride et al. show that more neurotic individuals are less likely to violate cyber-security policies [42]. Weirich and Sasse reported that people low on self-images and with self-admitted paranoia were more probable to not disclose personal information [75], thus showing a low level of social engineering susceptibility. They attribute this to fear of being held responsible to security breaches. Bansa et al. also report findings that neuroticism makes more sensitive towards privacy [5].

In the research reviewed, there were also examined some more specific personality traits that cannot be assigned to a single Big 5 trait. In Workman's study, high continuance commitment (significantly related to Openness, Conscientiousness, Extraversion, and Neuroticism [20]) has been shown to increase social engineering vulnerability [77]. He also showed that obedience to authority increases social engineering susceptibility, which supports Cialdini's principle of authority [77]. Unfortunately, to date there has been no study that showed a distinct relation between specific personality traits and obedience to authority. There exist other studies that investigate relations between personal attributes and susceptibility to the principles of influence (see [77] for examples). However, as these do not clearly relate to comprehensive personality theories, they have not been considered for the literature review in this work. The reason for this are discussions and criticisms that the discriminant validity of narrow personality traits is not sufficiently high [38]. Furthermore, it is questionable whether domain-specific personality traits like those used in some IT security studies can be considered personality traits at all, since one of the main aspects of personality traits is their manifestation in behaviour across different situations and contexts (see chapter 4).

5.2 Detailed SE-Personality-Framework and Research Proposals

The basic social-engineering-personality framework is very broad. Thus, it will be difficult to extract proposals for individual prevention measures. Furthermore, to date, only relations between personality traits and social engineering in general have been examined. Therefore, it is advisable to extend the framework to include specific relations between personality traits and single principles of influence. Thus, it will be possible to custom-tailor prevention measures not only for specific personality types but for different kinds of social engineering attacks as well. Figure 5.2 shows the proposed detailed

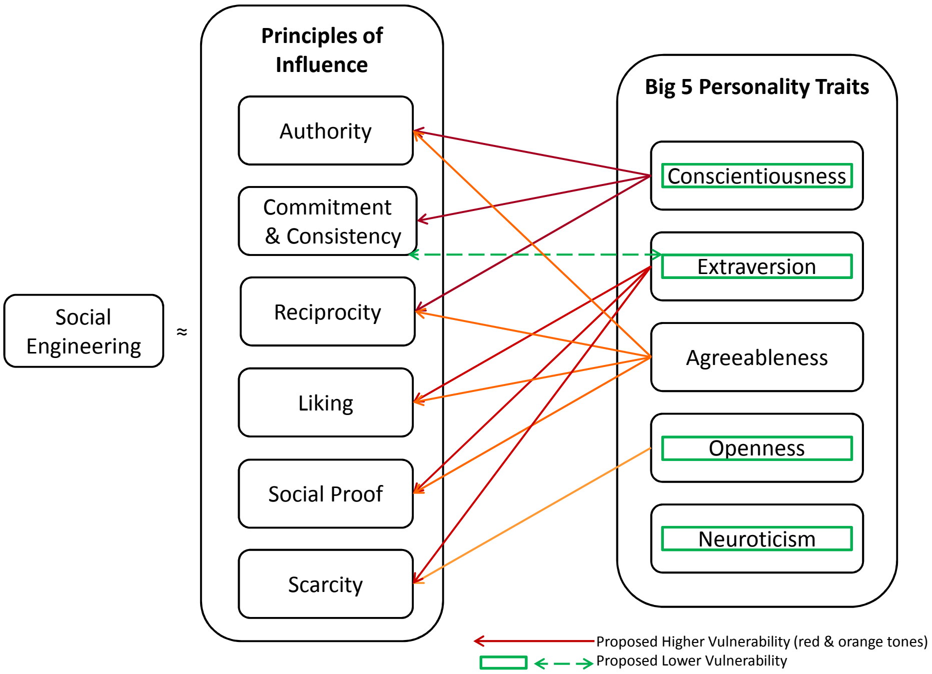

Figure 5.2: Detailed Proposed SE-Personality-Framework

framework. Below, for every personality trait, it will be described which relations are proposed. For some traits, general assumptions that have been found in the literature, are presented before the proposed relations. Afterwards, research proposals will be made how this framework could be validated.

Conscientiousness. Since conscientious people are known to strictly adhere to existing rules, sometimes even when common sense would question them, it is probable that they are more vulnerable to social engineering techniques that exploit rules. Thus, it is proposed that conscientiousness increases vulnerability towards the principles authority, reciprocity, and commitment-consistency. Commitment and consistency are supposed to only increase vulnerability when commitments are made publicly or refer to commitments concerning rules. For principles that do not exploit rules, such as liking, social proof, and scarcity, no relation or even a lower vulnerability due to conscientiousness is expected. Otherwise, conscientiousness can decrease social engineering susceptibility for every principle if there exists a useful security policy that contains a behavioural codex for coping with social engineering attacks. Awareness training should also prove essentially beneficial for conscientious individuals.

Extraversion. Liking and social proof should work especially well on extraverted individuals, since they rely on social aspects, and extraversion relates to sociability. The excitement seeking aspect of extraversion could lead to a greater vulnerability for the scarcity principle, since getting something scarce is usually described as exciting. However, high values in extraversion can also decrease the vulnerability towards commitment and consistency techniques, since extraverted individuals tend to have a lower preference for consistency.

Agreeableness. Individuals who were more trusting raised fewer concerns about privacy invasion by location based services [38], which could be assumed to be generalizable to fewer privacy concerns in agreeable individuals. This could then lead to a higher social engineering vulnerability in that those people would be more likely to disclose private information to an attacker. Regarding the principles of influence, an increased vulnerability towards authority, reciprocity, liking, and social proof is to be expected. Generally, every technique that involves other people and their opinions should work better on agreeable individuals, due to their trusting nature, their helpfulness, and their belief in the goodness of mankind.

Openness. Considering the lures used in social engineering attacks, openness to experiences and strong fantasy could lead to greater social engineering susceptibility [51]. On the other hand, openness has been associated with technological experience and computer proficiency [78]. Through this relation, openness could reduce social engineering vulnerability with respect to computer-based attacks, as more experience might lead to a better recognition of social engineering attacks. Of the principles, only scarcity is proposed to show a significant relation to openness, since it poses a perceived constriction of freedom – something aversive for an open individual. For the other principles, no relation to openness is expected.

Neuroticism. Parrish et al. propose that computer anxiety, which is associated with neuroticism, may protect the individual in regards to phishing or other computer-based social engineering attacks [51], mediated by greater caution when using a computer. Generally, it is proposed that neuroticism does not increase vulnerability towards attacks that use one of the principles of influence. Rather, neuroticism should act as a barrier against such attacks, since neurotic individuals often assume the worst from others in any situation.

This proposed framework can help researchers by giving them a structure of how personality relates to social engineering attacks, mediated by the principles of influence. However, the proposed relations still have to be validated by further research, especially the proposed links from personality traits to principles of influence. To validate this framework, social engineering attacks that each use only or basically one of the principles would need to be created. Ideally, for each principle one attack should be tried in realistic settings on multiple test subjects whose personality traits are measured before or after the tests. However, there are some problems associated with such a procedure. A realistic setting would ideally include some organization(s) that agree(s) to have social engineering attacks being tried on employees in the context of a penetration test. Problems with social engineering penetration tests will be discussed in the next chapter. Conducting six social engineering attacks per test person for a comprehensive and comparable data set could also cause problems, namely within-subject effects due to changes in perception of attacks and subsequently different behaviour in later tests compared to if that test would have been the first one. Another point of notice is possible stress and strain on test subjects that have to be told multiple times that they fell prey to a social engineering attack (again, see next chapter for a discussion). The obvious solution is to only conduct one social engineering attack per person. This results in a much greater logistic effort, since the number of test subjects multiplies along with the numbers of personality tests that have to be conducted and analysed. Furthermore, it has to be ensured that there are sufficient test subjects for each social engineering test, so that enough different personality profiles per test are available. Otherwise, no reliable statistical analysis can be conducted. The application of personality tests in organizations could also cause a problem, depending on the land of execution. As mentioned in chapter 4, in Germany, the conduction of comprehensive personality tests can be problematic in an organizational context. In this case, use of a business-oriented test like the BIP could be an alternative. Another possibility is not to use individual personality profiles but job profiles that depict the average personality profile of a job group. For the BIP, many of these job group profiles are available, and are based on a broad data base of German employees. The plus is that no individual personality tests have to be conducted, saving not only legal problems but also time and money. The downside is that the explanatory power is much lower with group averages than with individual characteristics. An alternative setting could be to conduct studies in a university context. Due to the complexity of the set-up, it would probably be wise to achieve a cooperation of several universities to ensure that enough test subjects can be recruited.

6 Prevention

According to the CSI survey [58], after a security incident, 42% of respondents provided additional security awareness training to end users, and 40,6% changed the security policies of the organization. These are the measures that are most often proposed in the reviewed literature, as well (see, for example, [62, 66, 76]). The crucial point is to convince all employees that everyone is vulnerable to social engineering attacks, and that the consequences and losses of such an attack can be substantial. Bezuidenhout et al. argue, that "individuals make themselves even more vulnerable to social engineering attacks by not expecting to ever be a victim of such an attack, and many will never know that they were a victim of such an attack" [6]. In not expecting to be the target of social engineering attacks, individuals may be more willing to disclose seemingly unimportant information. The argument that "it is often the person who thinks he is most secure who poses the biggest vulnerability" [15] points into a similar direction. People who think that they will recognize social engineering attacks effortlessly tend to ignore that the attacker is not only skilled at manipulating people but also uses situational variables for his benefit, like stressful environments where decisions must be made fast [6].

The problem with social engineering is that there is no universal defence against social engineering, since the behaviour that a social engineer elicits is usually socially desirable, and should therefore not be prohibited in general [66]. For defending against such attacks, it is essential to be able to analyse and interpret social relations and contexts reliably. In order to achieve this for every employee, organizations have to not only establish usable security policies but also have to create a culture of security that stimulates behavioural and attitude change.

In the next paragraphs, some examples for prevention techniques taken from the existing literature will be presented and discussed.

SEDA. Hoeschele and Rogers propose a prevention technique for social engineering attacks using the telephone, which completely circumvents the employee [34]. The Social Engineering Defense Architecture (SEDA), presented in Hoeschele and Rogers paper as a proof-of-concept, is supposed to detect social engineering attacks automatically by analysing phone conversations in real time, and thus determining if the caller is deceiving the callee. A logging facility for attack signatures, which enables criminological analysis of incidents, is also included in SEDA. The system is supposed to operate unobtrusively in the background, such that employees do not need to spend time on security checks. The voice signatures of every caller are stored in a personal information database alongside the name, corporate association, job title, and phone numbers used to place calls [34]. Thus, the system can also be used for authentication in phone calls, thereby increasing the difficulty for social engineers to pose as an insider, even if they have the knowledge and skill to pass as one. While SEDA does not address individual differences of the targeted employees, it nevertheless could support the employees in detecting social engineering attacks, and thus raise their knowledge about such attacks, if it ever becomes a business application. However, it could also lead to carelessness and lower awareness, if the employees rely on the system to detect attacks solely. Studies in realistic environments over a prolonged time period would be needed to evaluate this issue.

SEADM. Another prevention technique is the Social Engineering Attack Detection Model (SEADM), presented by Bezuidenhout et al. [6]. They argue that training as the predominantly preventive measure against social engineering is ineffective since the employees tend to forget the lessons learned soon after its execution. SEADM is proposed to mitigate this effect by providing employees with a simple decision tree that breaks down the complexity of a possible social engineering attack into binary questions that the employee can answer in a few seconds. While the idea is intriguing, the implementation raises a few questions. The employees are called to evaluate their emotional stability in the situation. When they do not feel emotionally stable, they should escalate the situation to another colleague or to their supervisor. This approach is escapist, and would in reality result in fast dismissal of an employee that frequently uses this functionality, since the employee would be seen as not attending to his duty. Furthermore, the misuse of this functionality is potentially high. The decision questions are based on the assumption that employees are able to look through the attackers disguise, otherwise they would not feel uneasy. This is an unrealistic assumption as well, since a well prepared and skilled social engineer will make the employee feel at ease most of the time. Since this approach is additionally not considering individual differences in employees, it will not be considered in the comprehensive prevention approach later in this chapter.

Workman's study on social engineering interventions. Workman investigated which of the interventions suggested by previous research were most beneficial for social engineering prevention [76]. He shows that punishment is a good deterrent for people who perceive greater fear, while social engineering prevention training is the best intervention for people who have higher levels of commitment and trust. Ethics training does not create an impact on the behaviour of targeted employees in social engineering attacks. Using a developmental approach that provides the test subjects with skills and a script improves social engineering security behaviour. However, none of the interventions changed the perceptions of threat vulnerability or severity, which was discussed above to be a crucial point. Otherwise, Workman argues that the test subjects in his study perceived social engineering as posing a significant and present danger regardless of the interventions, which somehow contradicts Bezuidenhout et al.'s argument, but could also be a side effect of the sample selection. Workman argues that, "while training did not affect behavior relative to perceptions of severity and likelihood of social engineering threats, when people perceive a social engineering threat as severe and likely, they are already more inclined to be cautious about lures" [76]. He nevertheless argues that, as vigilance sinks over time, there is a need to keep security awareness up when no threat is encountered in a certain time span.

The following three sections about penetration tests, security awareness training, and security-aware cultures present a comprehensive approach to preferably preventing and otherwise minimizing the impact of social engineering attacks. All three aspects should be considered when creating prevention measures against social engineering, since a better protection is provided when all three are applied.

6.1 Penetration Tests using Social Engineering Techniques

Penetration tests, which are in this context “applications of automated network vulnerability scanners to an operational site” [45], are commissioned by organizations, which want to test their security measures. They can include social engineering techniques. These can not only help the organization to recognize their vulnerabilities but can also help to raise awareness on social engineering issues. When a penetration tester performs social engineering attacks, there are some rules or requirements that should be followed to prevent problems with the organization, individuals, or the law.

Dimkov et al. [18] declare five requirements that should be satisfied such that the penetration test will be useful for the commissioning organization: the test should be realistic, with employees acting normally. All employees need to be treated respectfully. The test needs to be reliable, repeatable, and reportable, so all actions during the test should be logged, and the outcome should be prepared so that meaningful documentation of results and recommendations is available. These requirements are not without conflicts. For example, a realistic social engineering test relies on deception, which violates the respect requirement, and could hurt the employees’ trust in the organization and colleagues as well as produce stress and loss of productivity. Another example refers to reliability: when arbitrary employees are targeted, reliability is not warranted. More generally, reliability can be questioned in every setting including human behaviour, since it is not wholly predictable. Thus, a usable attunement of these requirements is needed. Dimkov et al. propose two methodologies, one environment-focused and the other custodian-focused, which in their opinion balance the requirements as best as possible while minimizing the impact of the tests on employees [18]. The methodologies are shown to be able to detect two kinds of vulnerabilities: “errors in implementation of procedural and physical policies by employees and lack of defined security policies from management” [18]. However, these methodologies limit the informational value about how well the organization is protected against real social engineering attacks. If an organization is interested in realistic settings, it should not use pre-proposed methods but decide together with the penetration tester, which scenarios and kinds of actions should be included in the tests.

From an organizational point of view, every action and attack scenario that is admissible should be written down into the contract between organization and penetration tester or into a separate document, called rules of engagement. Whatever document is used, it has to be signed by the management of the organization, and preferably as well by the security management and human resources management. If a works council or union exists, it is better to involve them as well. It has to be considered that less restrictive rules of engagement enable more realistic penetration tests but also violates the respect requirement more than restrictive rules. When attacks are conducted, the management and the security management should be informed in advance, so that in the case of attack detection and escalation, no law enforcement will be involved.

The individual employees that are targeted by the penetration tester with social engineering attacks also pose challenges that need to be addressed. Basically, this requires an ethics discussion. In medicine, social sciences, and psychology, there is a long tradition of ethics discussions and guidelines (see, for example, [24, 21, 23, 39, 56, 57]). The question that lies at the heart of these arguments is: what kind of actions are permissible from an ethical point of view, so that no subject of research (or, in the case of penetration testing, employee) is harmed, neither physically nor psychologically. Finn specifically discusses which conditions have to be fulfilled to justify the use of deception in research: the assessment cannot be performed without the use of deception; the knowledge obtained from the assessment has important value; the test involves only minimal risk (not more than the physical or psychological harm, which is normally encountered in daily live) and does not violate the rights and the welfare of the individual; the subjects are debriefed with relevant information after participation [23] (based on an aggregation by [18]).

On CSO [1], there are multiple articles that give advice on ethical use of social engineering. The recommendations in this paragraph are adapted from two articles written by Goodchild [28, 27]. She recommends to communicate to the commissioning organization that in the case of a first-time social engineering penetration test, it is to be expected that many employees fall for the attacks. These mistakes from a security point of view should be seen as an opportunity for learning, not for embarrassment or punishment. The benefit of these tests is to help do better in real incidents and future penetration tests. Embarrassing the organization or individual employees has to be avoided absolutely. This can be achieved by communicating the results of a penetration test in an audit as part of employee education. Generally, in an audit, no individual names should be publicly called. It is more useful and respectful to the individuals to debrief them privately. For example, when they fall prey to a phishing attack, the click on the link should open an educational page about what just happened, and how to avoid it in future. If the contract of the penetration tester also includes tries to getting access to assets, for example to a database, it is advisable to separate social engineering attacks on individuals from this part of the tests. After proving that getting access to the system via social engineering works, the exploitation should be set up as a separate phase of the project, where the penetration tester gets access from a collaborator inside. This approach does not involve social engineering but shows what can be exploited from a typical computer of an employee. It is not only safer, since no mail with malicious links or attachments can accidentally be forwarded to people outside of the scope of the test. It is also potentially less harmful for the employee, since no individual employee can be made responsible for the tester's successful exploitation of the system.

[1] `http://www.csoonline.com/topic/587703/social-engineering` (last called on July 7th, 2013)

The last point that needs to be considered is the law. There are many different laws that may apply to such a penetration test, ranging from data protection laws to civil and criminal law as well as from local to federal laws. It is absolutely necessary to ensure which laws apply for not getting in conflict with them. In most countries, it is illegal to pose as state officials. This cannot be bypassed. However, it is also illegal in most countries to impersonate a person within an organization. If this attack vector is included in the contract, it has to be assured in the contract that it will be allowed by the company, and will not be prosecuted. In this case, it is usually better not to impersonate a real person but to develop a fictional employee.

For computer-based social engineering, there exists an open source tool-kit, the Social Engineering Tool-kit (SET). It was created by David Kennedy, and integrates different tools that are designed to perform advanced social engineering attacks [31, 53]. This tool-kit enables penetration testers to easily construct computer-based attacks, which leaves more time for the direct interactions.

The recommendations given above can be subsumed into a few simple steps that should be followed to ensure the success of a penetration test that includes social engineering:

1. Set up a detailed contract containing what the commissioning organization wants to be tested, including attack vectors as well as single actions.
2. If some of the actions the contract contains violate local law, make sure that the contract specifications allow this in the context of the test.
3. Make sure that the employees concerned are not harmed, neither physically not psychologically. Particularly, stress should be kept at a minimum, and privacy should not be violated.
4. Separate social engineering attacks from system intrusion.
5. Inform the management and security management prior to attacks, such that no awkward situations arise.
6. Debrief the concerned employees without embarrassing them. Communicate mistakes as opportunities for learning and improvement.
7. Document the actions and results carefully.
8. Provide meaningful documentation of results and recommendations.

6.2 Security Awareness Training

Security awareness training should, as the name suggests, raise the awareness of all people inside the organization to security relevant topics. Although it is the most commonly used prevention measure against social engineering incidents, its impact, namely continuously higher security awareness and a higher detection rate as well as defence rate against social engineering, is controversial. Especially long term effects seem to be rather hard to obtain. In one paper [15], a penetration tester recounts that as many as 40% of employees who passed extensive awareness training still fell for a phishing attack conducted as measurement of success for the training. Although it is widely known that no 100% security can be achieved, in this case, most of the lacking impact of awareness trainings can be attributed to poor training design and content. There are a wide range of scientific publications available on how to construct good trainings with lasting effects, touching entire sciences, theories and topics like social sciences, learning psychology, cognitive theory, mental model construction, intrinsic motivation, transfer climate construct or organizational development to name only some of them (see, for example, [4, 8, 25, 40, 68]. Due to the complexity of this topic, it is beyond the scope of this work to discuss it further.

The most critical point to achieve in security training in general and particularly in social engineering prevention, is to get everyone in the organization to remain highly sensitive with respect to information dissemination. The basis of this are thorough and usable security and privacy policies, which are discussed in the next section. These policies need to be communicated, and their usage in daily routine trained during the awareness training. More precisely, policies can only make a difference, if the people who are supposed to adhere to them not only know their content but are also motivated and enabled to use them practically.

After having raised the awareness initially, it is crucial to keep it high. This can be achieved by making the message visible. For example, posters with simple, catchy, and relevant messages could be suspended in a prominent place. These posters should be exchanged regularly, so that attention for them will be kept high. Another possibility is to reward employees who show high levels of awareness, for example by putting messages like “Thank you for adhering to our policies” on desks without sensitive documents or password-stickers lying about after working hours. A regular competition for the best security awareness is also possible. It is important to keep the employees engaged, so feedback and suggestions should constantly be solicited. The American government has started the campaign “Stop. Think. Connect.” [2] as a global cyber-security awareness campaign. While it is intended to educate about online security, some of the free resources that are provided by the campaign could be used for awareness measures.

One main point mentioned by several authors is to make the people feel personal relevance for security related topics [15, 31, 59]. This does not only apply to social engineering but to IT security in general. The best way to achieve this is to connect security to their personal and private lives, because only if it matters to them personally, they can start to care about it in a work context as well. Show them what happens to their private computer if they open a malicious software or how keystroke loggers or trojans work. Hadnagy emphasizes not only the importance of private relevance but personal relevance in general [31]. He includes into training sessions colourful demonstrations that are intended to create a

[2] `http://stopthinkconnect.org/` (last called on July 7th, 2013)

greater impact. For example, he asks a participant at the start of the training to type a password, which the participant thinks to be safe, into Hadnagy's computer. While Hadnagy starts with the training, a password cracker is set to work. Usually, it cracks the password after a few minutes, which creates an immediate and drastic eye opener for the participants. This is a practical application of inoculation theory.

McGuire argues in his inoculation theory that previous exposure to a weakened persuasion attack makes individuals more resistant to further persuasion attempts, much like a vaccination with a weakened form of a disease creates resistance against the original disease [46]. It is assumed by McGuire that persuasion targets need to be supplied with the motivation to defend their attitudes, and the ability to do so effectively. This assumption is echoed in the call for more personal relevance of security awareness, and the allocation of adequate defence mechanisms through policies and training. A related topic is the effect of forewarning: Cialdini and Petty show that only forewarning of a persuasive intent of a message reliably causes resistance, while forewarning of the content of said message causes either resistance or acquiescence to the message [12].

Based on inoculation theory and forewarning, Sagarin et al. conducted three studies to examine the impact of a treatment that was designed to produce resistance to deceptive persuasion messages [60]. In the first study, they were able to show that after the resistance treatment, illegitimate persuasive appeals became less persuasive, while legitimate persuasive appeals became more persuasive. This was shown to be generalizable to other contexts and preservable over time in the second study. In the third study, the participants showed beliefs of invulnerability towards deceptive persuasion. These were dispelled in the study, which maximized the resistance against deceptive persuasion. The studies showed that two factors contribute to resistance to illegitimate appeals: perceived undue manipulative intent of the enquirer and perceived personal vulnerability towards manipulation. One important result of these three studies is that people frequently consider other people as being vulnerable to illegitimate manipulation, while they see themselves as more or less immune against such persuasion attempts. This illusion of immunity leads to weaker resistance against illegitimate persuasion attempts. If this illusion is dispelled, resistance can be maximized. This finding is in line with a saying from security practitioners, who argue that "it is often the person who thinks he is most secure who poses the biggest vulnerability. Some experts believe executives are the easiest social engineering targets" [15].

Applying these results to security awareness training demonstrates that it is not enough to teach employees about attack vectors, vulnerabilities, and defence mechanisms, thus increasing their persuasion knowledge. Rather, after teaching the knowledge, thus making the employees confident that they can resist these attacks, it is essential to demonstrate to them their individual vulnerability towards these attacks [62]. Additional research is needed to create methods for demonstrating individual vulnerabilities towards each of the six principles of influence in an ethical and realizable manner [60]. To underline the need for individual trainings referring not only to the principles of influence but also to different personality traits, a study by McBride et al. shall be highlighted. They showed that individuals with different personality traits characteristics reacted differently to the same training scenarios [42]. Thus, they conclude that security education, training, and awareness (SETA) programs should be adopted to individual's personality in order to achieve maximum impact. Hirsh et al. support these results by showing that "adapting persuasive messages to the personality traits of the target audience can be an effective way of increasing the messages' impact, and highlight the potential value of personality-based communication strategies" [33]. This is another explanation why current one-size-fits-all security awareness trainings do not attain the desired impact.

6.3 Policies and Security-aware Cultures

As mentioned before, comprehensive and usable security and privacy policies are the basis for security awareness trainings. They need to explicitly include a taxonomy, which information can be given to whom in which circumstances. Particularly, a clear policy regarding non-disclosure of personal information like credentials and passwords has to be established. Regarding all seemingly non-critical information, it should always be stated in the policy that the employee has to question whether the enquirer really needs to know the requested information. Moreover, a check on the authenticity of an enquirer, for example by callback or in-person verification, should be included in the policy. A strict compartmentalization of roles and competencies as well as allocating information on a need-to-know-basis seems to be a sensible security measure. However, it is contrary to many modern organizational theories that propose an organic structure and open communication [76]. The organization has to decide, whether this should be included in the security policies. While a detailed policy is desirable from a security point of view, it has to be kept in mind that the policies need to stay usable. When they become too cluttered, the employees will not bother to adhere to every detail. These parts of policies just mentioned could be laid down in strict rules that employees need to follow, which would be appropriate, for example, in a call centre setting. Another approach is to enable employees to identify an impostor, not according to strict rules but by sensitising the employees to the attributes of social engineering attacks and tactics used, for example, the principles of influence that have been discussed in this work, including the special vulnerabilities caused by personality traits.

However, simply creating and publishing these policies is not enough. The policies have to be continually kept alive by creating and maintaining a security-aware organizational culture. This culture is an essential part of prevention measures. It does, however, not relate to individual differences, which is the main focus of this work. Therefore, only a short overview will be given. There are many publications available, which discuss organizational cultures in general and particularly security-aware cultures (see, for example [17, 52, 63, 68, 70]). Weßelmann [74] emphasizes that especially in smaller organizations, the management is the main orientation for employees concerning security behaviour. They have to address the topic, and create sufficiently flexible procedures that support the employees in making better heuristics-based decisions. Another important part of the organizational culture is the communication style. It needs to enable informal possibilities to contact other colleagues and supervisors, whenever a request seems suspicious. Moreover, when a superior is behaving non-compliant concerning to security policies, employees have to be enabled to contradict him without fear of discipline. Failure management needs to be integrated into the culture. This means that it has to be accepted as normal that failures happen, and are not punished severely, but instead are used as learning opportunities. Clearly defined procedures and accountabilities for reporting suspicious incidents should also be established [76].

7 Conclusion and Outlook

The aim of this work was to examine how individual factors, and specifically personality traits relate to success or failure of social engineering attacks. Based on existent research, the Social-Engineering-Personality-Framework has been developed, which proposes distinct relations between each of the personality traits of the Big 5 Theory and the six principles of influence, which are underlying principles of social engineering attacks. Thus, differences in vulnerability to social engineering can be explained. This framework can guide future research in providing specific hypotheses about these relations. Since this thesis is a theoretical work, the more specific relations that are proposed need to be evaluated in future research. The present work helps to explain why so many social engineering attacks are successful, and proposes a comprehensive prevention plan that integrates penetration tests using social engineering with security awareness trainings and the development of an ongoing security-aware organizational culture. Relating to the context of existing social engineering research, the Social-Engineering-Personality-Framework can complement Tetri et al.'s Social Engineering Framework (as described in chapter 2.2, see figure 2.1). The latter accentuates that social engineering is not limited on the previously overemphasized interaction between attacker and targeted employee but also includes situational factors like policies and the targeted employee's interpretation of the situation. While it is argued in this work that these factors are influential, and should not be overlooked, Tetri et al.'s framework carries the risk of marginalizing the interaction between attacker and targeted employee. The suggested Social-Engineering-Personality-Framework addresses a specific relation between the two actors, namely, how the target's personality affects the success of the social engineering attack, thus complementing the previous framework regarding to an important and hitherto mostly unattended aspect. A future publication could integrate both frameworks with regard to content and graphical presentation. A practical application of the present work can be achieved, when the proposed relations of the framework have been evaluated: specific tailor-made prevention measures can be created, which account for these relations, and use them to decrease individual vulnerability to social engineering attacks.

Acknowledgement

I would like to thank my supervisors, Professor Gollmann and Sven Übelacker, for enabling me to write about such an interesting and prevailing topic. Without the support from both, this work would not have been possible.

Bibliography

[1] F. W. Abagnale and S. Redding. Catch me if you can, 1980.

[2] S. Abraham and I. Chengalur-Smith. An overview of social engineering malware: Trends, tactics, and implications. *Technology in Society*, 32(3):183 – 196, 2010. doi:10.1016/j.techsoc.2010.07.001.

[3] J. Asendorpf. *Psychologie der Persönlichkeit*. Springer DE, 2004.

[4] T. T. Baldwin and J. K. Ford. Transfer of training: A review and directions for future research. *Personnel Psychology*, 41(1):63–105, 1988.

[5] G. Bansal, F. Zahedi, and D. Gefen. The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. *Decision Support Systems*, 49(2):138 – 150, 2010. doi:10.1016/j.dss.2010.01.010.

[6] M. Bezuidenhout, F. Mouton, and H.S. Venter. Social engineering attack detection model: SEADM. In *Information Security for South Africa (ISSA), 2010*, pages 1–8, 2010. doi:10.1109/ISSA.2010.5588500.

[7] J. W. Brehm. A theory of psychological reactance. *New York*, 1966.

[8] R. H. Bruning, G. J. Schraw, and R. R. Ronning. *Cognitive psychology and instruction*. ERIC, 1999.

[9] L. McM. Bujold. *Diplomatic immunity*, volume 14. Baen Books, 2002.

[10] R. B. Cialdini. *Influence: science and practice, 5th edition*. Pearson, 2009.

[11] R. B. Cialdini and N. J. Goldstein. Social influence: Compliance and conformity. *Annu. Rev. Psychol.*, 55:591–621, 2004.

[12] R. B. Cialdini and R. E. Petty. Anticipatory opinion effects. *Cognitive responses in persuasion*, pages 217–235, 1981.

[13] R. B. Cialdini, M. R. Trost, and J. T. Newsom. Preference for consistency: The development of a valid measure and the discovery of surprising behavioral implications. *Journal of Personality and Social Psychology*, 69:318–318, 1995.

[14] R. B. Cialdini, W. Wosinska, D. W. Barrett, J. Butner, and M. Gornik-Durose. Compliance with a request in two cultures: The differential influence of social proof and commitment/consistency on

collectivists and individualists. *Personality and Social Psychology Bulletin*, 25(10):1242–1253, 1999.

[15] CSO Online. The Ultimate Guide to Social Engineering. Accessed: 2013-06-19. URL: `http://www.csoonline.com/article/701042/cso-s-ultimate-guide-to-social-engineering`.

[16] A. Darwish, A.E. Zarka, and F. Aloul. Towards understanding phishing victims' profile. In *Computer Systems and Industrial Informatics (ICCSII), 2012 International Conference on*, pages 1–5, 2012. `doi:10.1109/ICCSII.2012.6454454`.

[17] D. R. Denison. *Corporate culture and organizational effectiveness.* John Wiley & Sons, 1990.

[18] T. Dimkov, A. van Cleeff, W. Pieters, and P. Hartel. Two methodologies for physical penetration testing using social engineering. In *Proceedings of the 26th Annual Computer Security Applications Conference*, pages 399–408. ACM, 2010.

[19] W. Durant. The story of philosophy: The lives and opinions of the world's greatest philosophers author: Will durant, publisher: Pock. 1991.

[20] J. Erdheim, M. Wang, and M. J. Zickar. Linking the Big Five personality constructs to organizational commitment. *Personality and Individual Differences*, 41(5):959–970, 2006.

[21] A. Felnhofer, O. D. Kothgassner, and B. U. Stetina. Cyberethics. Ethik im Kontext der Online Forschung. *Ethik in der Psychologie*, pages 181–192, 2011.

[22] L. Festinger and J. M. Carlsmith. Cognitive consequences of forced compliance. 1959.

[23] P. R. Finn. Research Ethics: Cases and Materials, chapter The ethics of deception in research, 1995.

[24] National Commission for the Proptection of Human Subjects of Biomedical and MD. Behavioral Research, Bethesda. *The Belmont report: Ethical principles and guidelines for the protection of human subjects of research.*

[25] J. K. Ford et al. *Improving Training Effectiveness in Work Organizations. Series in Applied Psychology.* ERIC, 1997.

[26] M. Friestad and P. Wright. The persuasion knowledge model: How people cope with persuasion attempts. *Journal of consumer research*, pages 1–31, 1994.

[27] J. Goodchild. 3 tips for using the Social Engineering Toolkit. Accessed: 2013-07-06. URL: `http://www.csoonline.com/article/705106/3-tips-for-using-the-social-engineering-toolkit`.

[28] J. Goodchild. Social engineering in penetration tests: 6 tips for ethical (and legal) use. Accessed: 2013-07-06. URL: `http://www.csoonline.com/article/732251/social-engineering-in-penetration-tests-6-tips-for-ethical-and-legal-use`.

[29] D. Gragg. A multi-level defense against social engineering. *SANS Reading Room, March*, 13, 2003.

[30] R. Guadagno and R. B. Cialdini. Online persuasion and compliance: Social influence on the Internet and beyond. *The social net: Human behavior in cyberspace*, pages 91–113, 2005.

[31] C. Hadnagy. *Social engineering: The art of human hacking*. Wiley, 2010.

[32] N. Hammele. *Rechtliche Kriterien und Grenzen bei der Erhebung und Verwendung personenbezogener Daten im Rahmen von Arbeitsverhältnissen*. GRIN Verlag, 2011.

[33] J. B. Hirsh, S. K. Kang, and G. V. Bodenhausen. Personalized Persuasion Tailoring Persuasive Appeals to Recipients' Personality Traits. *Psychological science*, 23(6):578–581, 2012.

[34] M. Hoeschele and M. Rogers. Detecting Social Engineering. In Mark Pollitt and Sujeet Shenoi, editors, *Advances in Digital Forensics*, volume 194 of *IFIP — The International Federation for Information Processing*, pages 67–77. Springer US, 2005. `doi:10.1007/0-387-31163-7_6`.

[35] R. Hossiep and M. Paschen. *Das Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung: BIP*. Hogrefe, Verlag für Psychologie, 2003.

[36] M. Huber, S. Kowalski, M. Nohlberg, and S. Tjoa. Towards Automating Social Engineering Using Social Networking Sites. In *Computational Science and Engineering, 2009. CSE '09. International Conference on*, volume 3, pages 117–124, 2009. `doi:10.1109/CSE.2009.205`.

[37] L.J. Janczewski and L. Fu. Social engineering-based attacks: Model and new zealand perspective. In *Computer Science and Information Technology (IMCSIT), Proceedings of the 2010 International Multiconference on*, pages 847–853, 2010.

[38] I. Junglas and C. Spitzmuller. Personality Traits and Privacy Perceptions: An Empirical Study in the Context of Location-Based Services. In *Mobile Business, 2006. ICMB '06. International Conference on*, pages 36–36, 2006. `doi:10.1109/ICMB.2006.40`.

[39] G. P. Koocher and P. Keith-Spiegel. *Ethics in psychology: Professional standards and cases*, volume 3. Oxford University Press, 1998.

[40] D. H. Lim and M. L. Morris. Influence of trainee characteristics, instructional satisfaction, and organizational climate on perceived learning and training transfer. *Human Resource Development Quarterly*, 17(1):85–115, 2006.

[41] J. Littman. *The Fugitive Game: Online with Kevin Mitnick: The Inside Story of the Great Cyberchase*. Little, Brown & Co. Inc., 1996.

[42] M. McBride, L. Carter, and M. Warkentin. Exploring the Role of Individual Employee Characteristics and Personality on Employee Compliance with Cybersecurity Policies. 2012.

[43] R. R. McCrae and O. P. John. An Introduction to the Five-Factor Model and Its Applications. *Journal of Personality*, 60(2):175–215, 1992. doi:10.1111/j.1467-6494.1992.tb00970.x.

[44] Robert R McCrae, Paul T Costa, Jr, and Thomas A Martin. The NEO–PI–3: A more readable revised NEO personality inventory. *Journal of Personality Assessment*, 84(3):261–270, 2005.

[45] J. P. McDermott. Attack net penetration testing. In *Proceedings of the 2000 workshop on New security paradigms*, pages 15–21. ACM, 2001.

[46] W. J. McGuire. Inducing resistance to persuasion: Some contemporary approaches. *Advances in Experimental Social Psychology*, 1:191–229, 1964.

[47] S. Milgram. Some conditions of obedience and disobedience to authority. *Human relations*, 18(1):57–76, 1965.

[48] K. D. Mitnick and W. L. Simon. *The art of deception: Controlling the human element of security*. Wiley, 2002.

[49] J. G. Mohebzada, A. El Zarka, A. H. Bhojani, and A. Darwish. Phishing in a university community: Two large scale phishing experiments. In *Innovations in Information Technology (IIT), 2012 International Conference on*, pages 249–254. IEEE, 2012.

[50] G. L. Orgill, G. W. Romney, M. G. Bailey, and P. M. Orgill. The urgency for effective user privacy-education to counter social engineering attacks on secure computer systems. In *Proceedings of the 5th conference on Information technology education*, CITC5 '04, pages 177–181, New York, NY, USA, 2004. ACM. doi:10.1145/1029533.1029577.

[51] J. L. Parrish Jr, J. L. Bailey, and J. F. Courtney. A Personality Based Model for Determining Susceptibility to Phishing Attacks. *Little Rock: University of Arkansas*, 2009.

[52] C. Paulsen and T. Coulson. Beyond Awareness: Using Business Intelligence to Create a Culture of Information Security. *Communications of the IIMA*, 11(3):35–54, 2011.

[53] N. Pavkovic and L. Perkov. Social engineering toolkit - a systematic approach to social engineering. In *MIPRO, 2011 Proceedings of the 34th International Convention*, pages 1485–1489, 2011.

[54] B. Pellens. Theoretische und empirische Forschungsmethoden im Rahmen von Abschlussarbeiten. Accessed: 2013-07-14. URL: http://www.iur.ruhr-uni-bochum.de/imperia/md/content/iur/homepage/lehre/forschungsmethoden.pdf.

[55] T. R. Peltier. Social engineering: concepts and solutions. *Information Systems Security*, 15(5):13–21, 2006.

[56] U. Rauchfleisch. *Nach bestem Wissen und Gewissen: die ethische Verantwortung in Psychologie und Psychotherapie*. Verlag für Medizin. Psychologie, 1982.

[57] S. J. Reiser, A. J. Dyck, W. J. Curran, et al. *Ethics in medicine: historical perspectives and contemporary concerns*. MIT press Cambridge, MA:, 1977.

[58] R. Richardson. 2010/2011 CSI Computer Security Crime and Security Survey, 2011. Accessed: 2013-06-19. URL: `https://cours.etsmtl.ca/log619/documents/divers/CSIsurvey2010.pdf`.

[59] T. Ryan and G. Mauch. Getting in Bed with Robin Sage. In *Black Hat Conference*, 2010.

[60] B. J. Sagarin, R. B. Cialdini, W. E. Rice, and S. B. Serna. Dispelling the illusion of invulnerability: The motivations and mechanisms of resistance to persuasion. *Journal of Personality and Social Psychology*, 83(3):526–541, 2002.

[61] J. F. Salgado. The Big Five Personality Dimensions and Counterproductive Behaviors. *International Journal of Selection and Assessment*, 10(1-2):117–125, 2002. `doi:10.1111/1468-2389.00198`.

[62] J. W. Scheeres. Establishing the human firewall: reducing an individual's vulnerability to social engineering attacks. Technical report, DTIC Document, 2008.

[63] E. H. Schein. *Organizational culture and leadership*, volume 356. Wiley. com, 2006.

[64] B. Schneier. *Secrets Lies: Digital Security in a Networked World*. Wiley Publishing, Inc., 2004.

[65] B. Schneier. The Psychology of Security. In S. Vaudenay, editor, *Progress in Cryptology – AFRICACRYPT 2008*, volume 5023 of *Lecture Notes in Computer Science*, pages 50–79. Springer Berlin Heidelberg, 2008. `doi:10.1007/978-3-540-68164-9_5`.

[66] S. Schumacher. Die psychologischen Grundlagen des Social Engineerings. *Magdeburger Journal zur Sicherheitsforschung*, 1:1–26, 2011.

[67] J. Shropshire, M. Warkentin, A.C. Johnston, and M.B. Schmidt. Personality and IT security: An application of the five-factor model. In *Proceedings of the Americas Conference on Information Systems*, pages 3443–3449, 2006.

[68] M. T. Siponen. A conceptual foundation for organizational information security awareness. *Information Management & Computer Security*, 8(1):31–41, 2000.

[69] P. Tetri and J. Vuorinen. Dissecting social engineering. *Behaviour Information Technology*, 2013. `doi:10.1080/0144929X.2013.763860`.

[70] S. Übelacker. IT-Sicherheit, Unternehmenskulturen und wirtschaftsbedrohende Kriminalität. diploma thesis, University of Ulm, 2002.

[71] Verizon RISK Team. 2012 Data Breach Investigations Report, 2012. Accessed: 2013-06-19. URL: `http://www.verizonenterprise.com/resources/reports/rp_data-breach-investigations-report-2012-ebk_en_xg.pdf`.

[72] A. Vishwanath, V. Herath, R. Chen, J. Wang, and H. Raghav Rao. Why do people get phished? testing individual differences in phishing vulnerability within an integrated, information processing model. *Decision Support Systems*, 51(3):576 – 586, 2011. `doi:10.1016/j.dss.2011.03.002`.

[73] W. Voigt. *Wie ich Hauptmann von Köpenick wurde: mein Lebensbild*. J. Püttmann, 1909.

[74] B. Weßelmann. Maßnahmen gegen Social Engineering. *Datenschutz und Datensicherheit - DuD*, 32(9):601–604, 2008. `doi:10.1007/s11623-008-0143-3`.

[75] D. Weirich and M. A. Sasse. Pretty good persuasion: a first step towards effective password security in the real world. In *Proceedings of the 2001 workshop on New security paradigms*, pages 137–143. ACM, 2001.

[76] M. Workman. A test of interventions for security threats from social engineering. *Information Management & Computer Security*, 16(5):463–483, 2008.

[77] M. Workman. Wisecrackers: A theory-grounded investigation of phishing and pretext social engineering threats to information security. *Journal of the American Society for Information Science and Technology*, 59(4):662–674, 2008. `doi:10.1002/asi.20779`.

[78] A. B. Woszczynski, P. L. Roth, and A. H. Segars. Exploring the theoretical foundations of playfulness in computer interactions. *Computers in Human Behavior*, 18(4):369–388, 2002.