Received: 8 October 2015 | Revised: 12 May 2016 | Accepted: 19 April 2017

DOI: 10.1002/jip.1482

WILEY

RESEARCH ARTICLE

On the anatomy of social engineering attacks—A literature-based dissection of successful attacks

Jan-Willem Hendrik Bullée[1] | **Lorena Montoya**[1] | **Wolter Pieters**[2] | **Marianne Junger**[3] | **Pieter Hartel**[4]

[1]Services, Cyber-security, and Safety Group (SCS), Faculty of EEMCS, University of Twente, PO Box 217 Enschede, 7500 AE, The Netherlands
[2]Faculty of Technology, Policy and Management, Delft University of Technology P.O. Box 5015 2600 GA Delft
[3]Faculty of Faculteit Behavioural, Management and Social Sciences, University of Twente, P.O. Box 217, 7500 AE Enschede
[4]Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, P.O. Box 5031 2600 GA Delft

Correspondence
Jan-Willem Hendrik Bullée, Services, Cyber-security, and Safety Group (SCS), Faculty of EEMCS, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands.
Email: j.h.bullee@utwente.nl

Abstract

The aim of this study was to explore the extent to which persuasion principles are used in successful social engineering attacks. Seventy-four scenarios were extracted from 4 books on social engineering (written by social engineers) and analysed. Each scenario was split into attack steps, containing single interactions between offender and target. For each attack step, persuasion principles were identified. The main findings are that (a) persuasion principles are often used in social engineering attacks, (b) authority (1 of the 6 persuasion principles) is used considerably more often than others, and (c) single-principle attack steps occur more often than multiple-principle ones. The social engineers identified in the scenarios more often used persuasion principles compared to other social influences. The scenario analysis illustrates how to exploit the human element in security. The findings support the view that security mechanisms should include not only technical but also social countermeasures.

KEYWORDS

deception, information security, literature study, persuasion, social engineering

1 | INTRODUCTION

Social engineering is the art of exploiting the weakest link in information security systems (i.e., the people who use them; Bosworth, Kabay, & Whyne, 2014). The targets are deceived to release information or perform a malicious action on behalf of the offender (Huber, Kowalski, Nohlberg, & Tjoa, 2009). This paper focusses on the social influence

techniques offenders use in social engineering attacks to make their targets comply. Success stories from the social engineering literature were analysed to answer the research questions.

Information security incidents are often caused by human failure (Chan, Woon, & Kankanhalli, 2005), rather than technical failure (Schneier, 2000). Because humans handle information systems, information security needs to take not only the technical but also the human element into account. The attack on the human element of security is called "social engineering." This technique consists of using social influences to convince people that the offender (e.g., social engineer) is whom he claims or pretends to be. The offender takes advantage of people to obtain information or knowledge one should not have (Gupta, Agrawal, & Garg, 2011). Social engineering constitutes a security risk because it can be used to bypass intrusion detection systems, firewalls, and access control systems. One of the dangers of social engineering attacks is their harmless and legitimate appearance so that targets are unaware of being victimised (The Federal Bureau of Investigation, 2013; Hadnagy & Wilson, 2010). The result of a social engineering attack can be disastrous (e.g., crippled corporate networks, identity theft, or monetary loss; Gupta et al. ,2011).

Security experts have stated that as our culture becomes more dependent on information technology and technical prevention improves, social engineering will be the greatest threat to any security system Rouse (2006). Information security has traditionally been treated as a technical problem, resulting in information security teams being staffed solely by engineers (Rhee, Kim, & Ryu, 2009). This biased perspective of information security has led to the human factor being underestimated.

The field of human error, in a group or organisational context, has been widely researched (e.g., Chikudate, 2009; Edmondson, 1996; Zhao & Olivera, 2006). The type of human error that relates to unintentional errors is called a "slip." Slips refer to failures in executing the behaviour as planned and are mainly caused by attentional failure (Whittingham, 2004; Zhao & Olivera, 2006). Examples of slips are (a) a cashier forgetting to apply the discount to one product, (b) not properly closing a fire escape, or (c) leaving a USB key containing a list of confidential data on the train. The individual knows how to perform the task, has the intention to do it, but does not do it properly (Zhao & Olivera, 2006). Another type of error is that induced by external parties with the intention to make one fail a procedure (i.e., offenders using influencing techniques to manipulate their targets into compliance). Examples are (a) sharing credit card details with a stranger, (b) providing physical access to someone without appropriate credentials, or (c) supplying information that is needed for a spy to access the corporate network. The latter three examples relate to social engineering attacks.

Social engineering attacks can be explained using the analogy of the stack of Swiss cheese slices, known as the Swiss cheese model or the cumulative act effect (Reason, 1990). This model depicts the various elements of an organisation as layers. In an ideal world, these layers are intact; however, in the real world, they resemble slices of cheese with holes. A single layer that has a hole poses no problem. However, if there are holes in multiple layers and these align, an error can take place, that is, the system is penetrable, and therefore, an employee could become a victim of social engineering. Offenders can use influence techniques to create holes in the cheese or to align them.

This study investigates the errors induced by external parties, in the context of information security. Following the analogy of the Swiss cheese model, this study aims to identify the number, characteristics, and alignment of cheese slices (i.e., attack steps), which led to successful social engineering attacks (i.e., system failure).

From a psychological point of view, social engineering is part of decision making. Because people do not have the cognitive capacity to process all information, decision making involves using rules of thumb (i.e., heuristics; Cialdini, 2009). These mental shortcuts (resulting from experience and genetics) work well in most circumstances. However, when a heuristic goes wrong, a cognitive bias occurs (Gigerenzer, 1991; Tversky & Kahneman, 1974). Social engineers (i.e., the offenders) are well aware of the flaws in human logic and nudge the heuristics of their targets into systematic errors (i.e., cognitive biases) to make them comply (Bosworth et al., 2014; Dang, 2008; Kennedy, 2011; Luo, Brody, Seazzu, & Burd, 2011; Twitchell, 2009).

Studies on traditional crime have analysed both offenders and victims. However, in computer-related crime, it is difficult to find offenders for research purposes. In criminology, it is common to analyse the crime from the perspective of the targets themselves. Information extraction techniques such as interviews or questionnaires can be used to get details about the crime. On the other hand, it is harder to get the the offender's point of view, which is especially the case when the target is unaware of being victimised because the offender is unknown to him. This study aimed to

gain an insight into social engineering attacks from an offender's perspective, and the sources used are accounts by social engineers describing social engineering scenarios. The central question in this paper is how are social influences successfully used by social engineers to persuade their targets to comply with their requests?

A typical example of a social engineering attack from *The art of deception: Controlling the human element of security* (Mitnick & Simon, 2002):

> "Hi," says the voice at the other end of the line. "This is Tom at Parkhurst Travel. Your tickets to San Francisco are ready. Would you like us to deliver them, or would you like to pick them up?" "San Francisco?" says Peter. "I'm not going to San Francisco." "Is this Peter Abels?" "Yes, but I don't have any trips coming up." "Well," the caller says with a friendly laugh, "Are you sure you don't want to go to San Francisco?" "If you think you can talk my boss into it … " says Peter, playing along with the friendly conversation. "Sounds like a mix-up," the caller says. "On our system, we book travel arrangements under the employee number. Maybe somebody used the wrong number. What's your employee number?" Peter obligingly recites his number.

From this social engineering example, it is clear that the caller tricks the employee into disclosing his employee number. By having this small piece of specific information, one has knowledge that can be used to mimic an insider. By using a small lie, the caller made the employee fail to execute the "correct" behaviour (which in this case would be to callback the travel agency on a known number).

The first step in unravelling this complex topic is to find out how an individual can be persuaded to comply with malicious requests from offenders. By analysing offender accounts, we aim to find commonly used methods in social engineering attacks.

1.1 | Principles of persuasion

Once a person is a target, the offender can use social influences to change the odds of compliance in his favour. Compared to gullibility (i.e., the willingness to believe someone or something in the absence of reasonable proof; Greenspan, 2008), persuasion is a property of the offender, rather than of the target. A subset of six social influences (referred to as persuasion principles) were investigated by Cialdini (2009): authority, conformity, reciprocity, commitment, liking, and scarcity.

Authority is the principle that describes people's tendency to comply with the request of authoritative figures. If people are unable to make a thorough decision, the responsibility to do so is transferred to the group or someone they believe is in charge. Crisis and stress activate the behavioural trait of responsibility transition.

Conformity, or social proof, is the act of imitating the behaviour of other people. Members of the in-group have a stronger feeling of group safety compared with members of the out-group (Asch, 1951).

Reciprocity refers to the giving of something in return. The target feels indebted to the requester for making a gesture and even the smallest gift puts the requester in an advantageous position.

Commitment refers to the likelihood of sticking to a cause or idea after making a promise or adhesion. In general, when a promise is made, people will honour it, which increases the likelihood of compliance (Cialdini, 2009).

Liking someone puts that person in a favourable position. People tend to like others who are similar in terms of interests, attitudes, and beliefs.

Scarcity occurs when a product, service, or information has limited availability. People therefore perceive an increased value and attractiveness towards these products which makes them more desired than others.

1.2 | Effectiveness of persuasion principles

Persuasion principles have been researched for several decades, and there is empirical evidence that shows their effectiveness. Below, we present a short overview of meta-analyses related to each persuasion principle.

The most famous study that illustrated authority is the classical shock experiment performed by Stanley Milgram (1963). Sixty-six percent of the participants did not hesitate to deliver a deadly dose of 450 V to a human test

subject, which they were instructed to perform by a man they believed to have legitimate authority. An evaluation of 23 replication studies, conducted over the past 35 years, studied the obedience to authority paradigm. Almost half of the studies (i.e., 11 studies) showed a lower rate of obedience compared to the initial study by Milgram, ranging from 28% to 65%. The remaining 12 studies showed an equal or higher rate of obedience, ranging from 66% to 91% (Blass, 1999). Although the studies found different rates of obedience, it can be concluded that authority is an important behavioural phenomenon.

In 1965, Salomon Asch conducted an experiment on the effects of conformity. One of the outcomes was that 75% of the participants followed the majority, even when the majority stated an incorrect answer. This study was replicated many times in various contexts. A meta-analysis on conformity, containing 133 studies showed it to be effective (Bond & Smith, 1996). The level of conformity depends on the number of people (i.e., size of the majority) displaying a certain behaviour and can be estimated as $I = \frac{1}{e^{4n}}$ where $n = \frac{1}{e^{N^{1.75}}}$. In this function, I is the level of conformity, and N the majority size (Tanford & Penrod, 1984).

Reciprocity can be illustrated by means of reciprocal concessions, also known as the "door-in-the-face" technique (DitF). Compliance is the result of first rejecting the extreme initial request and then asking for a moderate alternative. The requester changes the initial request, whereas the receiver feels inclined to change as well, resulting in the change from a "no" to a "yes" (Cialdini et al. 1975). A meta-analysis ($k = 22$ studies, $N = 3,164$ subjects) showed that the DitF technique is effective. Those in the DitF group have higher odds of complying than those in the control group (Pascual & Guéguen, 2005).

Commitment is sometimes referred to as the "foot-in-the-door" (FitD) technique. Once a person has been induced to comply with a small request, he or she is more likely to comply with a larger demand (Freedman & Fraser, 1966). A meta-analysis ($k = 22, N = 3,124$) showed that the FitD technique is effective. Those in the FitD group have higher odds of complying than those in the control group (Pascual & Guéguen, 2005). The mechanisms behind DitF and FitD are diametrically opposed. A comparison ($k = 22, N = 3,192$) to find out which technique is more effective showed that both are equally effective (Pascual & Guéguen, 2005).

The effect of liking on self-disclosure has been investigated extensively, in particularly the question: "Do we disclose more to people we like?" A meta-analysis ($k = 31$) showed that people indeed have the tendency to disclose more personal information to people they like. Furthermore, there is little evidence that females and males differ in this respect (Collins & Miller, 1994).

Unlike the other principles, no meta-analysis was found for the scarcity principle; therefore, an individual study is discussed. Scarce products have an increased perceived value. The idea of not having something or a potential loss plays an important role in human decision making. Health researchers demonstrated this effect in a study involving breast cancer awareness pamphlets. The pamphlets stating the potential losses of not screening had significant more effect than pamphlets stating what the potential gain was (Meyerowitz & Chaiken, 1987).

The effect of scarcity was compared to conformity and a control condition, involving 153 university students in two emotional states: (a) fear and (b) romance. Results showed that (a) fear-state conformity appears to be more persuasive than the control condition (nonconformity), (b) fear-state scarcity appears to be less persuasive than the control condition (nonscarcity), (c) romance-state scarcity appears to be more persuasive than the control condition (nonscarcity), and (d) romance state conformity appears to be less persuasive than the control condition (nonconformity). Thus, fear improves the effect of conformity and romance improves the effect of scarcity (Griskevicius et al. 2009).

For a summary of meta-analytical findings, refer to Table 1. The p value indicates whether the principle was effective compared to the controls. This was the case for the conformity, reciprocity, commitment, and scarcity principles. For authority and liking, no such statistic was reported in the literature.

Based on Table 1, the persuasion principles were ranked on the basis of the compliance rate (i.e., effectiveness): (a) authority, (b) commitment, (c) reciprocity, and (d) conformity. Authority is therefore the persuasion principle with the strongest effect.

Persuasion principles seem to be effective, although none of the studies found focussed on the IT context. This research aims to answer the question: How are social influences successfully used by social engineers to persuade

TABLE 1 Overview of meta-analyses scores for each persuasion principle

Principle	*N*	*k*[e]	Compliance[d]	Statistic	*p*[a]	Reference
Authority	1,041[b]	23	62.5 [28–91]	–	–	(Blass, 1999)
Conformity	4,627	133	28.8 [2.1–60.1]	$d = .92$	.001	(Bond & Smith, 1996)
Reciprocity	3,164	22	41.1 [3.2–84.5]	$t(21) = 2.26$	.03	(Pascual & Guéguen, 2005)
Commitment	3,124	22	45.2 [3.2–100]	$t(21) = 2.71$	.01	(Pascual & Guéguen, 2005)
Liking	–	31	–	$d = .717$	.05	(Collins & Miller, 1994)
Scarcity[c]	311	1	–	$F(1, 305) = 5.34$	.021	(Griskevicius et al. 2009)
Reciprocity versus commitment	3,192	22	–	$t(21) = 0.1$	ns	(Pascual & Guéguen, 2005)

Note. – = Not Available; ns = not significant.

[a]The *p*-value tests whether the principle was effective in relation to the controls.

[b]*N* estimated by author of the present study based on available studies.

[c]Operationalised as romance.

[d]Average unweighed percentage of compliance [the number between brackets indicates the min and max].

[e]*k* indicates the number of studies included in the meta-analysis.

their targets to comply with their requests? One expectation was that in social engineering attacks, some persuasion principles would be used more often than others.

1.3 | Modality as moderator of persuasion

Social engineering is often associated with, but not limited to, calling a target and asking for a password (Winkler & Dealy, 1995). However, research on persuasion principles in the psychology literature shows that they are sometimes used over the telephone, but sometimes they are used face to face. An overview of differences in modality is presented below.

The influence of modality is illustrated in a meta-analysis on the DitF persuasion strategy. This study compared ($k = 56$) studies using the face-to-face modality with ($k = 31$) studies using the telephone modality. The combined result of $N = 7,641$ subjects, showed no statistical difference between the two (O'Keefe & Hale, 2001).

A meta-analysis on modality found no effect on compliance for both verbal ($k = 78, N = 7,138$) and behavioural ($k = 39, N = 3,125$) compliance (Feeley, Anker, & Aloe, 2012). Verbal compliance refers to saying that you will do something, whereas behavioural compliance to actual doing so. This means that modality does not "moderate" the relationship between strategy and compliance.

In a variation of the Milgram experiment (refer to Section 1.2), the authority (i.e., experimenter) left the room and gave instructions via the telephone. The results are an obedience rate of 20.5%, a significantly lower outcome compared to the 66% for the face-to-face condition (Milgram, 1965; 1974). The power of the authority seems to drop severely when there is no face-to-face contact.

The influence of modality was tested on a combination of product judgement (i.e., recall, attitude, and behavioural intentions) and type of goods (i.e., a product or a service) offered for sale (Szymanski, 2001). The product and the service were both presented via a face-to-face or telephone modality. The telephone modality had statistically significant higher scores (i.e., more influence) compared to the face-to-face modality. For goods, on the other hand, no difference was found in modality for behavioural intentions (Szymanski, 2001).

The results show that for some persuasion principles, modality moderates the relation between request and compliance, refer to Table 2. These mixed outcomes also indicate that there are other variables of influence and that the operationalisation is important.

TABLE 2 Overview of meta-analytical scores face to face/telephone

Principle	Moderator	*N*	k^b	p^a	Reference
Authority	–	80	1	<.001	(Milgram, 1965)
Reciprocity (DitF)	–	7,641	87	ns	(O'Keefe & Hale, 2001)
Reciprocity (DitF)	Verbal	7,138	78	ns	(Feeley et al. 2012)
Reciprocity (DitF)	Behavioural	3,125	39	ns	(Feeley et al. 2012)
Sales presentation	Goods	146	1	<.05	(Szymanski, 2001)
Sales presentation	Service	114	1	ns	(Szymanski, 2001)

Note. DitF = door-in-the-face; – = Not Available; ns= not significant.

[a]The *p* value tests whether the principle was effected by the modality.

[b]*k* indicates the number of studies included in the meta-analysis.

1.3.1 | Applying persuasion principles to bypass security

Research on persuasion principles is commonly associated with the social sciences (e.g., refer to Blass, (1999); Bond & Smith, 1996; Pascual & Guéguen, 2005; Collins & Miller, 1994). We next provide an example to illustrate how an offender could use each persuasion principle in an IT-related office environment setting:

Authority. An offender claims that he works in the IT department or that he is an executive in the company (Mitnick & Simon, 2002).

Conformity: The offender calls a target, says that he is conducting a survey, and names other people from the organisation who already participated. The survey embeds a series of questions that draw the victim into revealing the username and password.

Reciprocity: The offender calls a target and identifies himself as an employee from the IT department. The offender explains that there was a computer virus and that some data servers have to be restored. It would help the employee, if the target provides the password. The target is more likely to reciprocate because the caller offers him help.

Commitment: An employee is called from the finance department and asked for something small. The employee is called again later and asked for something bigger; this process is continued until the goal is reached.

Liking: An offender acts friendly, gives compliments, and states that he studied at the same university as the target hence emphasising similarities.

Scarcity: A call centre employee rings and gives a fake sales pitch. The target is not interested and wants to hang up the phone, but the offender rapidly mentions that the first 100 people who register on a given website will win two vouchers for a particular event

These six short scenarios are examples of how an offender can influence a target. The offender can use the six principles of persuasion to achieve a higher probability of compliance.

1.4 | Crime scripts

Understanding how offenders use social influences to convince their targets to comply is a key element in dissecting social engineering attacks. To dissect an attack into steps, we use the concept of the "crime script." Crime scripts relate to "attack templates," and in technical research areas, crime scripts are referred to as "attack life cycles" (Marconato, Kaaniche, & Nicomette, 2012). Cornish (1994) argued that crime scripts can be used to understand how a crime is executed. Crime scripts consist of sequential "script functions" and accompanying "script actions." These scripts organise the knowledge and understanding of the routine behavioural process: preconditions, initiation, actualisation, doing, and postconditions (Cornish, 1994). Breaking down crimes into small steps increases the comprehension of the crime as a whole. Each step in the crime script resembles a slice in the Swiss cheese model. However, this is not necessarily a one-to-one relationship because the Swiss cheese model relates to defender mechanisms; and it is possible that an offender bypasses several defence layers at once using a single attack step (or vice versa). Moreover, this dissection is useful for designing specific interventions.

Crime scripts have already been developed many times, for example, in the field of resale of stolen vehicles (Tremblay, Talon, & Hurley, 2001), organised crime (Rowe, Akman, Smith, & Tomison, 2012), illegal waste activity (Thompson & Chainey, 2011), drug manufacturing (Chiu, Leclerc, & Townsley, 2011), and serial sex offences (Beauregard, Proulx, Rossmo, Leclerc, & Allaire, 2007), whereas templates were used to identify rapists' target selection patterns (Beauregard, Rebocho, & Rossmo, 2010). However, to the best of our knowledge, there is no crime script analysis on social engineering attacks from the viewpoint of the offender. We therefore investigate the social influences used by an offender in a social engineering attack.

1.5 | Experimental context

Literature shows that social engineering works (Mann, 2008), that it is effective (Schellevis, 2011), and that targets are often unaware of being victimised (The Federal Bureau of Investigation, 2013; Hadnagy & Wilson, 2010). However, there is limited information available on the inner working of such attacks. Penetration testing reports occasionally surface, but these represent the proverbial needle in a haystack and are "uncontrolled" in the statistical sense. Books on social engineering (Hadnagy & Wilson, 2010; Mann, 2008; Mitnick & Simon, 2002; Mitnick, Simon, & Wozniak, 2011) provide an insight into how successful social engineering attacks are executed. To the best of our knowledge, no studies exist that identify social influences in social engineering attack scenarios.

Therefore, in this paper, social engineering cases are dissected to find out what social influences are used.

1.6 | Research question

The objective of the present research was to find out "How are social influences successfully used by social engineers to persuade their targets to comply with their requests?" Four questions and subquestions were formulated:

Q1. "What modality is used to execute social engineering attacks?"
Q2. "Which persuasion principles are used in the context of social engineering attacks?"
Q3. "How are the persuasion principles combined within a single step in social engineering attacks?"
Q3.1. "How many persuasion principles are used in an attack step?"
Q3.2. "How many attack steps are used in a crime script?"
Q3.3. "How are persuasion principles used in the course of a crime script?"
Q4. "What is the consistency among persuasion principles between two steps in the crime script?"
Q4.1. "Do principles differ over the course of a crime script?"
Q4.2. "How related are two consecutive attack steps within a crime script?"
Q4.3. "How related are the first and last step of a crime script?"

2 | METHOD

2.1 | Data selection

Goodreads is a social cataloging website where book lovers can review, rate, and catalogue what they have read. It currently contains 900 million books and 34 million reviews. There are over 100 books on Goodreads' bookshelf related to social engineering.[*] The following book inclusion criteria were used: (a) Social engineering involves a nontechnical social attack against the operator of a computer system, and (b) the book contains case studies illustrating the use of social engineering. The top four of books most often identified as social engineering by the community of Goodread members and met the two inclusion criteria were included in the analysis: (a) *The art of deception: Controlling the human*

[*] https://www.goodreads.com/shelf/show/social-engineering

element of security (Mitnick & Simon, 2002), (b) *Ghost in the Wires: My Adventures as the World's Most Wanted Hacker*" (Mitnick et al. 2011), (c) *Hacking the Human: Social Engineering Techniques and Security Countermeasures* (Mann, 2008), and (d) *Social Engineering: The Art of Human Hacking* (Hadnagy & Wilson, 2010).

According to Scopus†, the four books together are cited 240 times in scientific publications: 206, 8, 6, and 20 times. Compared to books written by other computer hackers, only *The Art of Deception* can be considered as highly cited (ie, 206 times), whereas *Social Engineering: The Art of Human Hacking* can be considered as moderately cited. The story of Richard Jones is told in *Underground: Tales of Hacking, Madness and Obsession on the Electronic Frontier*" (Dreyfus & Assange, 2012) and was cited 7 times. The memoires of Julian Assange, as in *Julian Assange: The Unauthorised Autobiography* (Assange, 2011) is cited 4 times. Former black hat hacker Kevin Poulsen wrote the book *Kingpin: How One Hacker Took Over the Billion-Dollar Cybercrime Underground* (Poulsen, 2011), which is cited 8 times.

In each book, in the analysis, social engineering is illustrated by means of success scenarios, executed by the authors, their friends, or other professionals from the field. The present analysis includes only scenarios that contain interactions between at least two humans. Scenarios that consist only of pure hacking, such as obtaining access to information via an open port on the server found using a port scanner were excluded because these lack human interaction. In addition, we excluded forms of threatening such as blackmailing and other human forms of interactions such as bribing, as we do not consider these to constitute social engineering. What is considered as social engineering is the use of deception and manipulation of the human element to release information to the offender (Bosworth et al. 2014; Dang, 2008; Huber et al. 2009). In total, there were 74 social engineering scenarios extracted for the analysis.

2.1.1 | Quality of the books and ecological validity (i.e., realism)

An internet search was performed to assess the quality of the four books and their ecological validity (i.e., their degree of realism). The ratings of the four books found in three on-line sources (i.e., books.google.com, amazon.com, and goodread.com) were analysed. The ratings show that more than 70% of the readers rated the books as being either good or very good. Some of the scenarios that originate from *The Art of Deception: Controlling the Human Element of Security*" were described as fictionalised (Mitnick & Simon, 2002). For the other three books, there was no evidence found about either fabrication or fictionalisation of stories, based on the book reviews at Goodreads and Amazon.

In an attempt to obtain more insight into the extent to which scenarios in *The Art of Deception* were fictionalised, the online legal research database Westlaw was queried. The results show that Kevin Mitnick was convicted two times: the first time in 1988 and the second time in 1996. In 1988, he was charged with "possession of unauthorised access devices" (in violation of 18 Code of Laws of the United States of America (U.S.C.) section 1029(a)(3)) and computer fraud (in violation of 18 U.S.C. section 1030(a)(4)). He pled guilty and was sentenced to 12 months in prison followed by a 3-year period of supervised release under special conditions. The company where the computer fraud was performed is mentioned in the book *The Art of Deception* (Mitnick & Simon, 2002); however, there are insufficient details in the court documents to validate the method (USA v. Mitnick, 1998). In February 1995, Mitnick was charged with "access device fraud" (in violation of 18 U.S.C. section 1029) and computer fraud (in violation of 18 U.S.C. section 1030). A 25 count indictment against him alleging "computer fraud," "wire fraud," and "interception of wire communications" relating to additional alleged illegal conduct by him was presented in September 1996. This indictment contains a list of organisations, which are mentioned in the scenarios used in *Ghost in the Wires* (Mitnick et al. 2011). Similarly, no details were found in the court documents to validate the method (USA v. Mitnick, 1996). The court documents obtained via the Westlaw online legal research service provided no additional insight into the scenarios. What was therefore found from this validation investigation was that some of the companies mentioned in the books' scenarios also appear in the indictments, therefore, it seems plausible that the book *The Art of Deception* is not a work of fiction.

† http://www.scopus.com

2.2 | Readers

All 74 scenarios were independently coded by two researchers. The first researcher (i.e., the first author) is a 29-year-old male PhD candidate with a background in both psychology and computer science. The second researcher is a 22-year-old female student assistant with a background in information and communication sciences. An interrater reliability analysis using the κ statistic was performed to determine the consistency among researchers.

2.3 | Procedure

To ensure agreement between the researchers, they both (a) processed a description of the persuasion principles, (b) performed coding on a test dataset of five scenarios, and (c) discussed the outcome of the training results. The description of the principles was the same as that of Section 1.1. The set of training scenarios was a random selection from all scenarios.

After the training, both readers agreed that analysing a scenario as a whole (refer to Figure 1) would bias the results, because multiple people can be involved and an offender can approach each individual differently. Therefore, it was decided to split the scenarios into attack steps, each containing a single interaction between two individuals. For example, if the offender first talked to employee *A* and next to employee *B* and finally to employee *C*, the scenario is split into three attack steps (refer to Figure 2). The persuasion principles used by the offender were coded for each attack step (refer to Figure 3). Figure 3 shows three interactions containing one, two, and two persuasion principles. All the

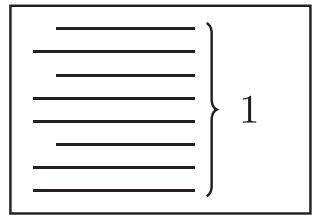

FIGURE 1 One scenario

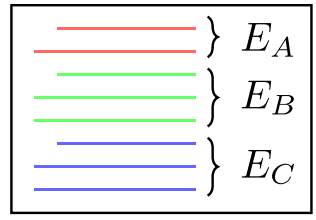

FIGURE 2 Three attack steps

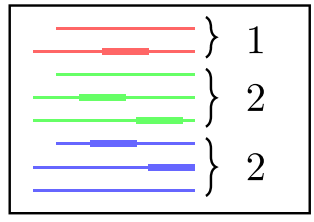

FIGURE 3 Five persuasion principles

scenarios were coded twice, meaning that the work was not split and concatenated afterwards but that instead, the resulting dataset consists of consensual results. After coding all attack steps, the interrater agreement was calculated. The scores of both readers were compared to generate the final dataset. If both readers identified the same principles for a given attack step, there was consensus. However, when there was a difference in the codes, the readers discussed the different views and came to a conclusion (the majority of differences related to one coder accidentally marking the wrong principle).

Figure 4 shows the dissection of a "real" scenario from Mitnick and Simon (2002). The scenario contains two attack steps, where each attack step contains one persuasion principle. In both attack steps 1 and 2, the offender uses impersonation together with the authority principle. In attack step 1, this was achieved by claiming to be an attorney, whereas in attack step 2, by claiming to be a staff member from the R&D department. In both attack steps, authority was operationalised by means of titles.

Cracker Robert Jorday had been regularly breaking into the computer networks of a global company. The company eventually recognized that someone was hacking into their terminal server, and could connect to any computer system at the company. To safeguard the corporate network, a dial-up password was required on every terminal server.

Robert *called*[a] the Network Operations Center `posing`[b] as an attorney with the **Legal Department**[c] and said he was having trouble connecting to the network. The network administrator explained that there had been some recent security issues, so all dial-up access users would need to obtain the monthly password from their manager. Robert wondered what method was being used to communicate each month's password to the managers and how he could obtain it.

It turned out that the password for the upcoming month was sent in a memo via office, mail to each company manager.

[a]Using the phone modality
[b]Impersonation
[c]Authority

Robert *called*[a] the company after the first of the month, and reached Janet, the secretary of a manager. He said, "*Janet, hi.* ***`This is Randy Goldstein`***[b] *in* ***Research and Development***[c]*. I know I probably got the memo with this month's password for logging into the terminal server from outside the company but I can't find it anywhere. Did you get your memo for this, month?*"

Yes, she said, she did get it.

He asked her if she would fax it to him, and she agreed. He gave her the fax number of the lobby receptionist in a different building on the company campus. He had already made arrangements for faxes to be held for him, and be forwarded to an on-line fax service. When this service receives a fax, the automated system sends it to the subscriber's email address.

The new password arrived at the email dead drop that Robert set up on a free email service in China. Best of all, he never had to show up physically at the location of the fax machine.

[a]Using the phone modality
[b]Impersonation
[c]Authority

FIGURE 4 Example: one scenario, two attack steps, and two persuasion principles

2.4 | Variables

There are two kinds of variables in the analysis, those related to (a) the crime script and (b) the attack step. The variables related to the crime script are "modality" and "steps." The variables related to the attack step are "persuasion principles" (six variables), "other," "first/last," and "former/latter."

The six dichotomous (persuasion principle) variables were dummy coded as 0 = *notused* and 1 = *used*. In case none of the six persuasion principles seemed appropriate, the variable other recorded other social influence tactics the offender used to deceive the target. The variable was a string variable hence allowing an open-ended response. In the open-ended responses, there was one suggestion given related to "overloading." Overloading can be used while conducting a questionnaire by putting the trick question between other questions. Due to the amount of information the brain has to process, there is a transition to a passive mental state, in which information is absorbed rather than evaluated (Janczewski & Colarik, 2008).

The modality variable was a string variable recording the modality used by the offender (e.g., face to face or telephone).

The steps variable was an integer recording the number of attack steps in each crime script (i.e., 1 means that there was a single attack step in the crime script).

The categorical variables first/last contained the persuasion principle(s) used in the first and last attack steps of a crime script.

The categorical variables former/latter contained the persuasion principle(s) used in a chronological attack step pair of a crime script.

2.5 | Analysis

The first question (i.e., What modality is used to execute social engineering attacks?) involves comparing the frequencies of different modalities to perform social engineering attacks. The variable modality was tested using cross tabulation and chi-square analysis.[‡]

The second question (i.e., Which persuasion principles are used in the context of social engineering attacks?) involves the frequencies of persuasion principles used. The variable persuasion principles was recoded into a single variable (i.e., an entry for each occurrence of the persuasion principle) and was tested using Fisher's exact test.[§]

The third question (i.e., How are the persuasion principles combined within a single step in social engineering attacks?) contains three subquestions: (a) "How many persuasion principles are used in an attack step?", (b) "How many attack steps are used in an crime script?", and (c) "How are persuasion principles used in the course of a crime script?" The aim of the question is to get an insight in how social engineering attacks are executed, which is relevant for developing countermeasures. The first subquestion involves the variable persuasion principles, which was recoded into a new variable containing the number of persuasion principles used for each attack step and it was tested using Fisher's exact test.[§]. The second subquestion compares path lengths using the steps variable, and it was tested using Fisher's exact test.[§] The third subquestion involves the variables persuasion principles and steps, and it was tested using Fisher's exact test.[§] In total, there were four analyses performed, one for each persuasion principle (except for scarcity and conformity due to insufficient observations).

For the fourth question, (i.e., What is the consistency among persuasion principles between two steps in the crime script?) contains three subquestions: (a) "Do principles differ over the course of a crime script?", (b) "How related are two consecutive attack steps within a crime script?", and (c) "How related are the first and last step of a crime script?" These subquestions enable a better understanding of the crime scripts. The first subquestion provides an insight into

[‡]The following two data assumptions must be met for chi-square analysis: (a) independence and (b) minimum frequency of five observations per cell (Field, Miles, & Field, 2012). Independence relates to putting a single observation in only one cell. In case one assumption is not met, the Fisher's exact test should be used instead.

[§]Fisher's exact test was used because the minimum number of observations was not met.

TABLE 3 Overview of variables and statistical test.

Question[a]	Variable[b]	Ideal test[c]	Assumptions		Applicable test[f]
			Ind[d]	Obs[e]	
Q1	Modality	chi-square	Y	N	Fisher's exact
Q2	Persuasion principles	chi-square	Y	N	Fisher's exact
Q3.1	Persuasion principles	chi-square	Y	N	Fisher's exact
Q3.2	Steps	chi-square	Y	N	Fisher's exact
Q3.3	Steps (authority)	chi-square	Y	N	Fisher's exact
Q3.3	Steps (commitment)	chi-square	Y	N	Fisher's exact
Q3.3	Steps (liking)	chi-square	Y	N	Fisher's exact
Q3.3	Steps (reciprocity)	chi-square	Y	N	Fisher's exact
Q4.1	Persuasion principles	chi-square	Y	Y	chi-square
Q4.2	Former/latter	chi-square	Y	N	Fisher's exact
Q4.3	First/last	chi-square	Y	N	Fisher's exact

[a]This column refers to the specific question and subquestion.

[b]This column refers to the variables that were used in the statistical test.

[c]Ideal test refers to which test ideally to use if all the assumptions were met.

[d]Ind refers to the independence assumption that in required for some tests.

[e]Obs refers to the minimal number of observations that is required for some tests.

[f]The applicable test given the assumption results.

the use of persuasion principles at different stages of the attack and involves the variables persuasion principle and steps. The variable persuasion principle was recoded as 0 = not using persuasion principle and 1 = using persuasion principle. This was tested using cross tabulation and chi-square. The second and third subquestions are used to analyse changes in tactics at a microattack level. The second subquestion compares two consecutive attack steps. It involves the variables former/latter, and it was tested using Fisher's exact test.[§] The third subquestion compares the first attack step with the last using the variables first/last. This was tested using Fisher's exact test.[§] Ultimately these analyses enable to find out whether the offender's principle selection is drawn at random or not, which would influence the design of awareness campaigns by security managers.

For an overview of all variables used in the different tests, refer to Table 3.

In this study, the focus is on the "execution" phase of the crime script, which follows the preparation and target selection phases. All the interactions (i.e., attack steps) found in the 74 scenarios were annotated with a time component to keep track of the chronological order.

3 | RESULTS

3.1 | Descriptive statistics

Seventy-four scenarios were analysed. This work resulted in the identification of 142 attack steps containing 180 occurrences of persuasion principles. Out of the 142 attack steps, 125 used persuasion principles whereas 17 involved other social influences. The number of attack steps containing persuasion principles is statistically different from 0, $\chi^2(1, N = 142) = 49.5, p < .001$. This means that there is a difference between the scenarios that use persuasion principles and the group that does not use them. The category of "other social influences" covers 12% of all social influences used by offenders in their attack steps. This category contains: (a) act ignorant 1× (5.9%), (b) creating curiosity 2× (11.8%), (c) distracting 1× (5.9%), (d) empathy/pity 2× (11.8%), (e) just ask for it 9× (52.9%), and (f) overloading 2× (11.8%).

TABLE 4 Combination of principles found in books

Principles	Where to find	Freq
OTHER - Comply	1:[47, 6, 16, 36, 30]	5×
OTHER - OTHER - Auth - Comply	1:[2]	1×
OTHER - Auth - Comply	1:[37, 11]	2×
OTHER - Auth - Auth - Auth Reci Comm - Auth - Comply	3:[5]	1×
OTHER - Auth Comm - Auth - Auth - Auth Reci Like - Comply	2:[4]	1×
OTHER - Auth - *Auth Reci* - OTHER - Comply	1:[27]	1×
OTHER - Like - Comply	2:[16], 4:[2]	2×
Auth - Comply	1:[5, 10, 12, 17, 21, 22, 25, 29, 35, 41, 45, 46], 2:[12, 14], 3:[3]	15×
Auth - Auth - Comply	1:[1, 26, 38, 44, 49], 2:[1, 11], 3:[1]	8×
Auth - Auth - Auth - Comply	1:[28,40]	2×
Auth - Auth - Auth - Auth - Auth - Comply	2:[18]	1×
Auth - Auth - *Auth Comm* - Comply	2:[13]	1×
Auth - Auth - *Auth Comm* - *Auth Comm* - OTHER - Auth - Comply	1:[34]	1×
Auth - Auth - *Auth Like* - Comply	2:[5]	1×
Auth - Auth - *Auth Reci* - Auth - Auth - *Auth Reci* - Comply	2:[8]	1×
Auth - Auth - Reci Comm Like - Comply	4:[1]	1×
Auth - *Auth Reci* - Comply	2:[10]	1×
Auth - Auth Reci Like - Comply	1:[14]	1×
Auth Comm - Comply	1:[8, 23] 2:[6]	3×
Auth Comm - Auth - Like - *Auth Like* - Comply	1:[3]	1×
Auth Comm Like - Comply	3:[2]	1×
Auth Comm Like - Auth - Auth - Comply	2:[7]	1×
Auth Comm Like - Auth - *Auth Reci* - Auth - Comply	2:[9]	1×
Auth Comm Like - Auth Comm Like - Comply	1:[9]	1×
Auth Comm Like Scar - Comply	2:[17]	1×
Auth Conf Comm Like - Auth - **Comm** - Auth - Auth - Comply	4:[3]	1×
Auth Like - Comply	1:[48], 2:[15], 3:[6]	3×
Auth Like - Auth - Comply	2:[3]	1×
Auth Reci - Comply	1:[7, 15, 24, 39], 2:[2], 3:[4]	6×
Auth Reci - Auth - Comply	1:[13]	1×
Auth Reci - **Conf** - Comply	1:[18]	1×
Auth Reci Like - Comply	1:[32]	1×
Like - Comply	1:[4, 20]	2×
Like - OTHER - Comply	1:[43]	1×
Reci - Comply	1:[31]	1×
Reci Comm Like - Comply	4:[4]	1×
	Total	74×

Note. The left-most principle constitutes the initial step in the attack.

Auth : In more than half of the attack steps, authority was used as single step.

Auth Reci : All 27 double principle scenarios contained authority.

Auth Reci Like : Liking is used in 91% of the triple principle scenarios.

Reci : Commitment, conformity, and reciprocity are only executed once as single principle.

OTHER : Seventeen occurrences being nonpersuasion principles

All attack steps (i.e., interactions) are summarised in Table 4, which shows where the scenario originated from. Table 4 also shows that (a) authority as a single principle is the most commonly used attack step; it is used in 76 (53.5%) of all attack steps, and it is identified in all six steps over time; (b) double principles (i.e., using two principles in one attack step) include authority in all 27 (100%) attack steps; (c) liking was used in 10 (91%) of the cases, which use three persuasion principles in an attack step; (d) both commitment, conformity, and reciprocity are only executed once as single principles; (e) the most frequently used attack path is (a) `Authority` - (b) `Authority - Compliance`. All crime scripts and attack steps listed on Table 4 are summarised in a single-tree structure (refer to Figure fig:Forest). The attack steps containing similar strategies (i.e., persuasion principles) at the same point in time are merged into a single node in the tree. This tree gives an overview of all the social engineering scenarios from the books and shows the frequent and less frequent attack paths. There are other studies using multiple crime scripts for generalisation purposes, for example, on the topic of wildlife trafficking (Lavorgna, 2014). However, this is the first study that combined multiple crime scripts into a single visualisation.

3.2 | Interrater agreement

The researchers' interrater reliability was $\kappa = .909$ ($p < .000$), 95% confidence interval [.874, .944], $N = 852$. The N represents 142 attack steps times six possible persuasion principles per step. The results indicate that there is an almost perfect agreement between the two researchers (Landis & Koch, 1977).

3.3 | Q1: "What modality is used to execute social engineering attacks?"

Out of the 142 attack steps, 123 (86.6%) were executed via telephone, and 18 (12.7%) were executed face to face and once via chat. There was a statistically significant difference in the modality used in executing persuasion principles in social engineering attack steps, $\chi^2(1, N = 142) = 43.983$, $p = .000$. Regarding modality of attack steps, the telephone modality was used considerably more often, but attack steps containing other means considerably less often.

3.4 | Q2: "Which persuasion principles are used in the context of social engineering attacks?"

Fisher's exact test showed that there was a significant difference in the use of principles during social engineering attacks, $p = .000$. The occurrence of the six persuasion principles is (a) 63% for authority, (b) between 10% and 13% for liking, reciprocity, and commitment, and (c) under 2% for scarcity and conformity (refer to Figure 5).

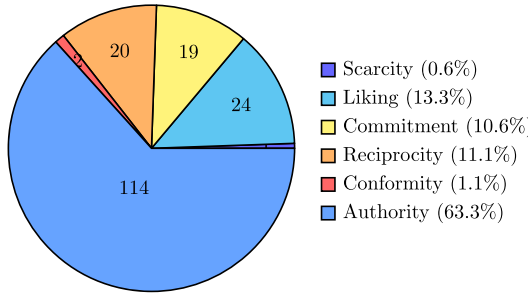

FIGURE 5 Persuasion principles used

3.5 | Q2: "How are the persuasion principles combined in social engineering attack steps?"

3.5.1 | Q3.1: "How many persuasion principles are used in an attack step?"

In total, 142 attack steps contain social influences; of which, 125 include persuasion principles, and the remaining 17 contain some other social influence.

Single attack steps contain up to four persuasion principles. The average number of persuasion principles used per attack step is $M = 1.44$ ($SD = .723$). There was a statistically significant difference in the number of occurrences of persuasion principles used in a single social engineering attack step, $p = .000$.

Regarding simultaneously used principles, single principles are used considerably more often whereas quadruple principles considerably less often, refer to Figure 6.

3.5.2 | Q3.2: "How many attack steps are used in an crime script?"

In total, 74 scenarios contain 142 attack steps. The shortest attack path contained a single step, whereas the longest attack path consisted of six attack steps. On average, the attack path has $M = 1.92$ ($SD = 1.311$) steps. There was a statistically significant difference in the number of attack steps used in crime scripts, $p = .000$. Regarding combining

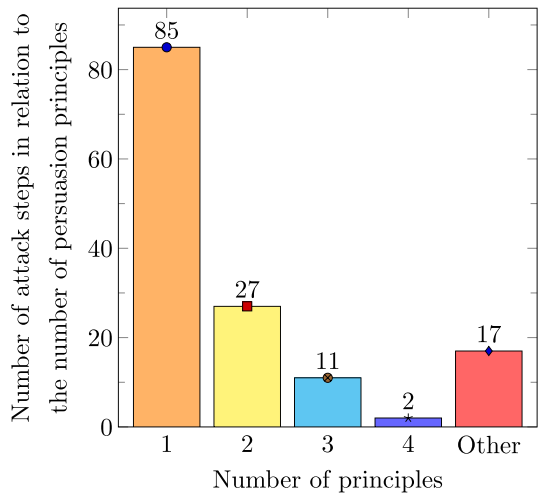

FIGURE 6 Number of principles used per interaction

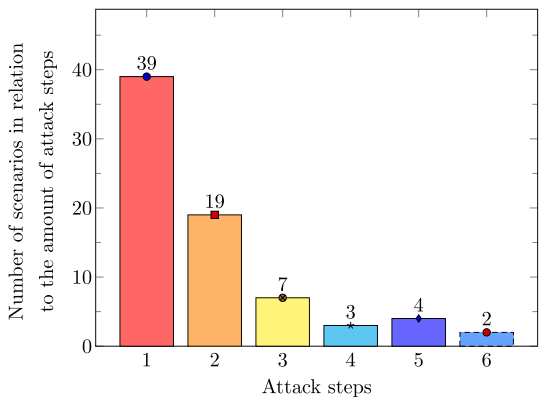

FIGURE 7 Number steps in an attack

attack steps, single step attacks are used considerably more often, but attacks containing six steps considerably less often, refer to Figure 7.

3.5.3 | Q3.3: "How are persuasion principles used in the course of a crime script?"

The combined principles were decomposed to obtain the occurrence of each principle (for absolute occurrences refer to Figure 8 and for relative occurrences to Figure 9). In the first step, before compliance, there were 101 occurrences of persuasion principles; in the second step, this decreased to 38, whereas in the sixth step before compliance, there were only two occurrences. There was a statistical significant difference in the use of authority ($p = .000$), liking ($p = .005$), commitment ($p = .003$), and reciprocity ($p = .002$) over time. Authority, liking, commitment, and reciprocity are used considerably more in the first step before compliance, but these are used considerably less in the sixth, fifth, and fourth

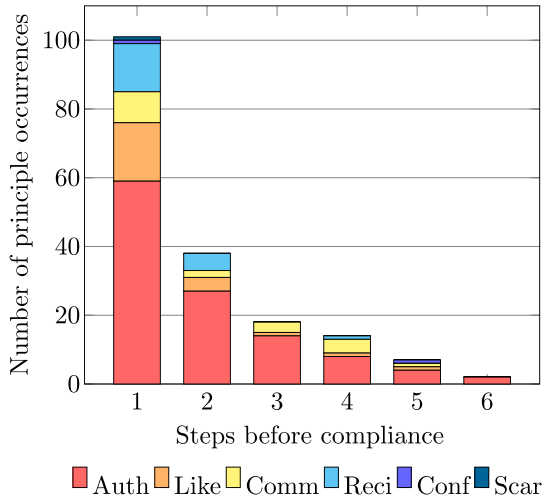

FIGURE 8 Absolute principle use over time

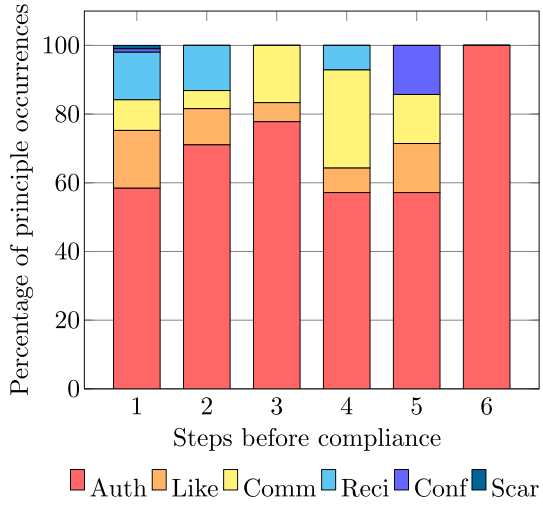

FIGURE 9 Relative principle use over time

step before compliance. Due to the limited occurrence of the scarcity and conformity principles, it was not possible to perform a statistical test.

3.6 | Q4: "What is the similarity between the different attack steps in the crime script?"

3.6.1 | Q4.1: "Do principles differ over the course of an attack?"

In a multiple step attack, the average number of steps is $M = 1.92$ ($SD = 1.311$) therefore, the last three attack steps before compliance comprise 88% of all attack steps. In the remaining 17 steps, there are eight single principle steps, which all contain authority; there are four double principle steps, which all combine authority with some other principle (i.e., commitment and reciprocity). Furthermore, there is one triple principle step consisting of authority, commitment, and liking, and there is one quadruple one containing authority, conformity, commitment, and liking. Finally, there are three steps containing another form of social influence than the six persuasion principles.

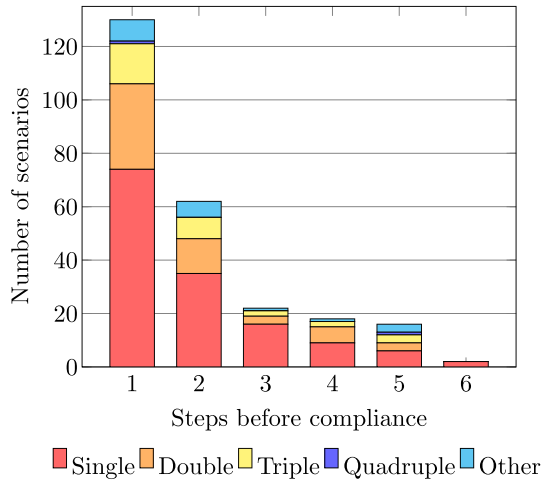

FIGURE 10 Absolute principle combination for each step before compliance

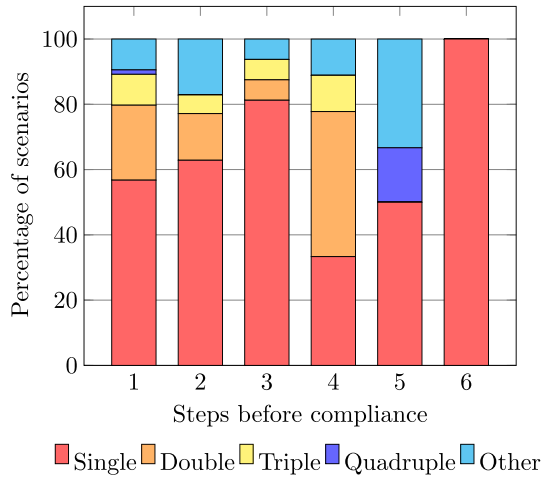

FIGURE 11 Relative principle combination for each step before compliance

TABLE 5 The number of combined principles in an attack step over time before the target complies with the request (N=142)

		Steps before compliance						
		1	2	3	4	5	6	Total
Principles	Single	42 (56.8%)	22 (62.9%)	13 (81.3%)	3 (33.3%)	3 (60%)	2 (100%)	85
	Double	17 (23.0%)	5 (14.3%)	1 (6.3%)	4 (44.4%)	–	–	27
	Triple	7 (9.5%)	2 (5.7%)	1 (6.3%)	1 (11.1%)	–	–	11
	Quadruple	1 (1.4%)	–	–	–	1 (20%)	–	2
	Other	*7 (9.5%)*	*6 (17.1%)*	*1 (6.3%)*	*1 (11.1%)*	*2 (40%)*	–	17
	Total	74 (100%)	35 (100%)	16 (100%)	9 (100%)	6 (100%)	2 (100%)	**142**

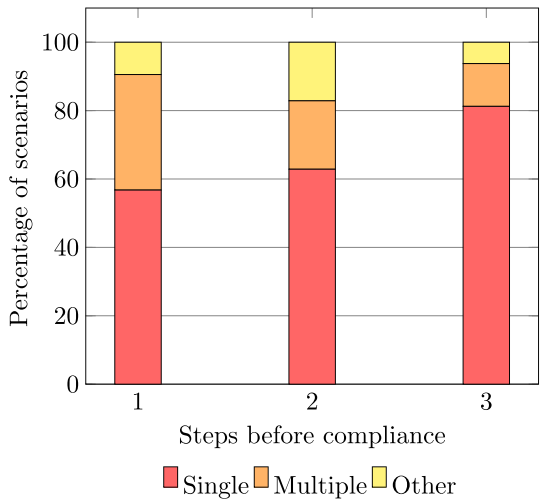

FIGURE 12 Use of principle combinations last three steps before compliance

Single principle attack steps are more popular than combined principle attacks (refer to Figure 6). Authority as a single step is used in 85.7% of the steps immediately before compliance, 90.9% of the second steps before compliance, and 91.7% of the third steps before compliance. The use of authority (i.e., in isolation) is relatively stable over time. The absolute use of principles over time is shown in Figure 10, and the relative use in Figure 11.

Despite the small numbers, in double principle attack steps the combination authority + commitment (3 times), liking (4 times), or reciprocity (7 times) are recurring combinations (refer to Table 4). In the final step before compliance, 67 attack steps used persuasion principles versus seven that used some other social influence. The number of scenarios containing persuasion principles in the last step before compliance is different from 0, $\chi^2(1, N = 74) = 29.108$, $p < .001$.

Table 5 shows an increase of attacks using multiple principles towards the end of the attack. Towards the moment of compliance, the relative use of single persuasion principles in an attack step drops from 81.3% to 56.8%, whereas the use of multiple persuasion principles in attack steps increased from 12.6% to 33.9%. This tendency is summarised in Figure 12.

3.6.2 | Q4.2: Comparison between two consecutive steps

In total, there are 68 consecutive attack steps identified in the 74 scenarios. The number of principles in the former attack step does not differ from that of its latter step, ($p = .768$).

There are 45 attack steps that contain persuasion principles; 39 (57.4%) attack steps began with authority, 26 (38.2%) steps succeeded with a single authority attack step whereas 10 (14.7%) of them consist of a combination of authority + commitment, reciprocity, or liking. Only three consecutive steps end with something other than authority. Finally, there are 13 steps that either start or end with a social influence other than a persuasion principle.

The combined principles in an attack step were further decomposed. One hundred fifteen consecutive individual principles were found in the 74 scenarios. There are 17 steps which either begin or end (i.e., succeed) with a social influence other than the six persuasion principles. The principles used in the former step do not differ with respect to those in the latter step ($p = .630$). Out of the remaining 99 consecutive persuasion principles, 74 (74.7%) consecutive principles begin with authority, whereas 49 (49.5%) of the consecutive principles also end with this principle. Furthermore, there are only 7 (7.07%) consecutive persuasion principles, which do not involve authority.

3.6.3 | Q4.3: Relation first and last step before compliance

There are 26 scenarios with more than one attack step. The number of persuasion principles in the first step does not differ from those in the final step, ($p = .515$).

The combined principles in an attack step were further decomposed. In total, there are 55 scenarios with individual persuasion principles. The number of persuasion principles used in the first attack step does not differ from those in the final step, ($p = .797$). There are 37 (66%) scenarios starting with authority, and it is used 24 times (42.9%) as the last step in the scenario. Only six (10.7%) scenarios do not involve this principle as either the first or last principle.

4 | DISCUSSION

In social engineering attacks, offenders use persuasion principles to change the odds of their target complying with their request. This study investigated how persuasion principles were used in successful social engineering attacks based on the accounts of social engineers.

The dissection of crime scripts shows that the anatomy of social engineering attacks consists of (a) persuasion principles (refer to Q2), (b) other social influences (refer to Q2), (c) deception, (d) real-time communication, and (e) telephone operation (refer to Q1). The heart of the social engineering attacks is shown in orange in Figure 13. This visualization contains the key elements of a social engineering attack because (a) approximately 80% of the crime scripts consist of one or two attack steps (refer to Figure 7); (b) approximately 80% of the attack steps consist of one or two persuasion principles (refer to Figure 6 and Table 5); (c) the most frequently used persuasion principles are authority and liking (refer to Figure 5); (d) besides persuasion principles, other social influences are used (refer to Q2); and (e) combined persuasion principles in attack steps always contain authority.

This study gives a unique insight into how offenders use social engineering successfully to perform their criminal act. In 88% of the attack steps, persuasion principles were used by the offender. Hence, in 12% of the attack steps, another social influence was used. We can therefore conclude that (based on their accounts) social engineers make frequent use of persuasion principles as social influences to make their targets comply with their request. Some principles are used more frequently than others. Given the similarity between the ranking of principles based on the success rates in the meta-analyses (refer to Table 1) and the ranking of principles in this study (refer to Figure 5), we believe that the occurrence of principles reflects their effectiveness in social engineering.

In order to draw conclusions about effectiveness, an experiment would need to be performed to verify the present conjectures. The experimental conditions can be controlled in an experiment (i.e., the use of the persuasion principles). Furthermore, the effectiveness can be determined because the number of subjects who comply and do not comply is known. When the experimental design and context are kept constant, the effectiveness of the persuasion principles can be calculated. Such an experiment could, for example, involve a so-called technical support scam. This involves a telephone fraud scam where the offender impersonates technical support service personnel (Arthur, 2010). The modus

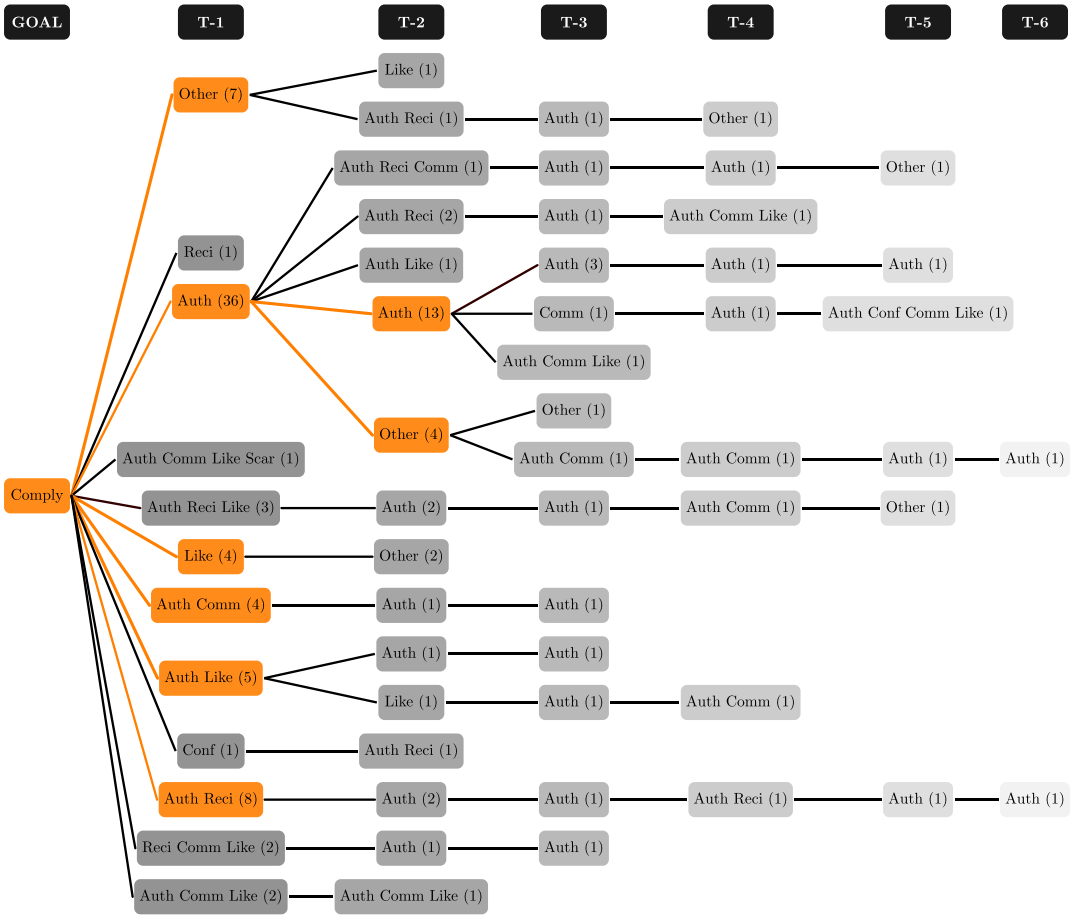

FIGURE 13 Tree structure of social engineering scenarios. The time flows from right (initial contact) to left (achieve goal). The frequency of occurrence is specified inside the brackets. The orange coloured nodes resemble the heart of social engineering attacks

operandi often includes the offender informing the target that there is a problem with their PC. To resolve the problem, the caller recommends buying a small software tool to prevent further damage. The use of individual or multiple persuasion principles can be included in the telephone script used by the offender. One instance of this experimental design is shown in (Bullée, Montoya, Junger, & Hartel 2016).

It was found that there are significantly more social engineering attacks executed via the telephone, compared to other modalities. We believe that this is because of the lower effort and risk that such an attack entails compared to one that involves being physically present. Physically going to meet the target implies additional risks: (a) the risk of being caught at the scene, (b) body language could hamper the plan, and (c) the attacker can only attack once as his face might have been recognized.

The outcome of an attack can be influenced if offenders are able to apply persuasion principles. Literature suggests that all principles can be effective. However, the success of principles depends on the context, operationalisation and final goal. Milgram (1965) showed that there was a significantly lower effect of authority when switching to the telephone modality. However, a "nurse experiment" showed a high rate of compliance when authority was applied over the telephone (Hofling, Brotzman, Dalrymple, Graves, & Pierce, 1966). This study shows that authority was the most frequently successful used principle. The reason it is most commonly executed over the telephone probably relates

to its relative low effort. Regarding authority, there are several factors that explain its effectiveness. One of the preconditions for authority is the institutional framework of modern society because childhood people are taught how to operate within an institutional system. This framework is initially in the form of regulated parental-adult authority, later through the authority of teachers in school and finally through a boss in a company or commander in the army (Milgram, 1963). Another precondition for authority is rewards for compliance to authority, whereas failure to comply results in a punishment. Authority (from a psychological point of view) is when a person is perceived to be in the position of social control for a particular situation (Milgram, 1963). The person claiming to be the authority will succeed if (a) someone expects an authority, (b) appropriate dress or equipment is used (e.g., lab coat, tag, and uniform), (c) there is absence of competing authority, and (d) there is absence of conspicuous factors (e.g., a 5-year-old child claiming to be a pilot). Unless contradictions in the information or anomalies appear, the authority will likely suffice. People respond to the appearance of authority, rather than the actual authority (Milgram, 1963). Another finding of this study is that only very few social engineering attacks begin or end with the scarcity principle; we assume that this reflects low success rates. Scarcity seems easy to operationalise, because this is a frequently used technique in sales and television advertisements (e.g., "only 25 products left" or "order **today** and get a 50% discount on the second item"). Furthermore, it seems that single step attacks only containing commitment, scarcity, or conformity are rare. The data shows that instead, Commitment and reciprocity are used in combination with authority. This could indicate that this combination of principles strengthens each other.

Knowledge about the principles, principle combinations and the time line in social engineering attacks is useful for designing countermeasures. The topic of countermeasures is discussed later in this section.

The use of a single principle occurs in almost 60% of the attack steps. The use of combined persuasion principles is used in less than 30% of the attack steps. This suggests that it is easier to operationalize an attack step consisting of a single principle compared to multiple principles. Although it is likely that by combining multiple principles in one step the effectiveness of that step increases, the operational complexity could outweigh its benefits. To the best of our knowledge, the issue of principle combination had not been discussed in the literature until now.

The results of this study show that the average number of attack steps (interactions) is two. This means that the crime script is short and that in order to make a target comply, only information/support from one other employee is needed. In the final step before compliance, the number of combined principles increases. This could indicate that to make the target comply, a boost is needed in the final attack step and that this is being achieved by combining principles in one attack step.

The results show no difference between the number of principles used in the first compared to the last attack step. Furthermore, there is no statistical difference in use of principles between the former and latter of two successive attack steps within a crime script. The results indicate that a single principle attack step is more likely to occur after a single principle attack step. Similarly, it is more likely that the use of authority is followed by authority. From this, we can conclude that offenders, like all other humans, might be creatures of habit in the sense that they stick to the method they initially chose. We assume offenders choose their method based on successful past attempts.

Moreover, regarding countermeasures to defeat social engineers, results showed that successful social engineering attacks most often use authority. However, because our society is built on the authority paradigm, it would be extremely difficult to counteract authority by itself. We believe that it is a better approach to use situational countermeasures. These could involve four mechanisms; (a) procedural, (b) environmental, (c) technical, and (d) behavioural.

Procedural countermeasures could, for example, involve the use a classification system for all organisational data, including employee and PC names, schedules, and software versions. Data above a certain classification threshold should not be allowed to leave the organisation. Furthermore, in the case of someone receiving a request, it should be determined if the request is legitimate (Mitnick & Simon, 2002). This can be done by verifying the identity of the requester (e.g., call-back policy or shared secret). An employee who receives a request for information should verify the identity of the requester and initiate the communication via a channel that is verified by the organisation. If someone claims to be a colleague, the employment status should be confirmed (e.g., lookup in employee directory or contact their manager for verification). After verifying the employment status, check the knowledge needs of that person

(e.g., check level of classification employee or contact his manager). It is important that the verification does not become a time-consuming process because employees might then ignore this.

Environmental countermeasures adjust the environment to encourage a desired behaviour. Research results show that social engineers are likely to operate via the telephone. One environmental countermeasure could involve placing stickers on telephones, to remind the user each time they use the telephone.

A technical solution could include using white and black lists. Telephone calls from numbers on the white list are connected, whereas those on the black list get terminated.

Behavioural countermeasures relate to adjusting the human: (a) by making employees aware that there are social engineers that use social influences to make people comply with their requests and they should realise that they are vulnerable (Muscanell, Guadagno, & Murphy, 2014), (b) by training people to spot a social engineering attack, (c) by making employees aware of why this is dangerous and what the implications are for the individual and the organisation, and (d) by distributing guidelines about what to do or not to if they are under a social engineering attack.

Muscanell et al. (2014) describe the best practices to resist social influences (i.e., persuasion principles). Best practices result in six questions to counteract the individual persuasion principles: (a) Authority: When approached by an authority, "Is this person truly whom he claims to be?" (b) Conformity: The fact that many others do something does not guarantee that it is a correct behaviour, hence "Would I do the same if I was alone in this situation?" (c) Reciprocity: "Why did I get this favour? Is this an act of kindness or part of a manipulation strategy?" (d) Commitment: Evaluate all events as independently as possible: "Do I really want this?" (e) Liking: Separate the request from the person: "What would I say if the request came from a different person?" (f) Scarcity: Once something is scarce, the perceived value increases: "Is this still an attractive offer if it wasn't scarce?"

The short-term effects of informing employees has been demonstrated by a group of students using social engineering to obtain office keys from university personnel. Those who received an information campaign complied significantly less frequently than those who were not informed (Bullée, Montoya, Pieters, Junger, & Hartel, 2015). A related relevant issue would be to assess how learning impacts compliance over time. We expect that "quick" interventions that are repeated on a regular basis are effective. Such an effect is already shown in the context of cardiopulmonary resuscitation (CPR) skill retention. In this study, the subjects were given a 4-min CPR training every 1, 2, and 3 months after the start of the study. The final result (after 6 months) was an increase from ±20% to ±70% of the subjects performing a perfect CPR (Sutton et al. 2011).

Finally, both the implementation and the effectiveness of the procedure should be tested. This can be achieved by trying to social engineer one's own employees in a controlled environment. If this is done regularly, the employees will most likely remain responsive and alert.

When conducting experiments on humans (e.g., employees) some ethical concerns must be taken into account. The Belmont report (1979) defines three ethical principles for the protection of humans during testing: (a) respect for persons, (b) beneficence, and (c) justice.

"Respect for persons incorporates at least two ethical convictions: first, that individuals should be treated as autonomous agents, and second, that persons with diminished autonomy are entitled to protection" (Belmont Report, 1979). Respect for persons means that people are free to participate or to decline participation in research. Furthermore, people who are not capable of making competent decisions by themselves should be guided by a capable guardian. Beneficence is defined as "persons are treated in an ethical manner not only by respecting their decisions and protecting them from harm, but also by making efforts to secure their well-being" (Belmont Report, 1979). This means that one should not harm the participants. Moreover, the possible benefits of participating should be maximised, and the potential harm should be minimised. "Who ought to receive the benefits of research and bear its burdens?" (Belmont Report, 1979) relates to the justice principle. Both the risks and benefits of the study should be equally distributed within the subjects.

One ethical challenge is the use of deception because it conflicts with the "respect for persons" principle. The use of deception might be acceptable if (a) the experiment does not involve more than minimum risk (i.e., harm or discomfort should not be greater than those experienced in daily life; Code of Federal Regulations, 2005), (b) the study could not

be performed without deception (subjects in laboratory studies may behave differently than they normally would or their behaviour may be altered because of the experimental setting), (c) the knowledge obtained from the study has important value, and (d) when appropriate, the subjects are provided with relevant information about the assessment afterwards (i.e., debriefing) (Finn & Jakobsson, 2007).

Dimkov, Pieters, and Hartel (2009) described the 5 R^* requirements in penetration testing research (which often uses social engineering): (a) realistic (the test should resemble a real-life scenario), (b) respectful (the test should be done ethically) (c) reliable (the test should not cause productivity loss of employees) (d) repeatable (repeating the test should result in similar results), and (e) reportable (all actions should be logged). Dimkov et al. (2009) identified conflicting requirements and noted that designing a penetration test involves finding a balance in the requirements. For a discussion of the three major ethics standpoints (i.e., [a] virtue ethics, [b] utilitarianism, and [c] deontology) refer to Mouton, Malan, Kimppa, and Venter (2015). Finally, it should be noted that in this study, the authors did not have any control regarding the ethical concerns in the scenarios as they are literature based.

4.1 | Limitations

The proposed study has four limitations: (a) Some of the scenarios by Kevin Mitnick that were used in the analysis might have been fictionalised to some extent. (b) The data set contained a limited number of observations. (c) The dataset could suffer from selection bias. It is possible that the authors could have favoured one scenario over another one. (d) The analysis in this research only contains success stories. This therefore provides a partial view and influences the conclusions that can be drawn. On the other hand, available data is limited and this analysis gives a unique insight into how offenders perform their offences.

Finally, we summarise these recommendations for future research. The details of all recommendations are already presented in Section 4; therefore, only a brief summary is presented in this section. First, the analysis of the four books shows that social engineering works. However, all scenarios in the books involved success stories; therefore, it is unclear what the success rate of a social engineering attack is. Experiments should be used to find the success rates. Second, one should investigate how persuasion principles influence each other when combined in an attack step to identify which ones have the likelihood of succeeding. Furthermore, a useful follow-up study could involve investigating if these social engineering attacks can be blocked or their effects reduced. Finally, it is possible that compliance depends on cultural aspects. The deployment of social engineering experiments in different countries could allow to identify cross-country differences.

ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 318003 (TREsPASS). This publication reflects only the author's views, and the Union is not liable for any use that may be made of the information contained herein. In addition, we would also like to thank Jessica Heijmans for her valuable contribution to this research.

REFERENCES

Arthur, C. (2010). Virus phone scam being run from call centres in India. (Newspaper Article). http://www.theguardian.com/world/2010/jul/18/phone-scam-india-call-cent&res.

Asch, S. E. (1951). *Effects of group pressure upon the modification and distortion of judgments*, Groups, Leadership, and Men (177–190). Oxford, UK: Carnegie Press .

Assange, J. (2011). *Julian Assange: The unauthorised autobiography*. Edinburgh: Canongate Books.

Beauregard, E., Proulx, J., Rossmo, K., Leclerc, B., & Allaire, J-F (2007). Script analysis of the hunting process of serial sex offenders. *Criminal Justice and Behavior*, *34*(8), 1069–1084.

Beauregard, E., Rebocho, M. F., & Rossmo, D. K. (2010). Target selection patterns in rape. *Journal of Investigative Psychology and Offender Profiling*, *7*(2), 137–152.

Belmont Report (1979). The belmont report : Ethical principles and guidelines for the protection of human subjects of research.

Blass, T. (1999). The Milgram paradigm after 35 years: Some things we now know about obedience to authority1. *Journal of Applied Social Psychology*, *29*(5), 955–978.

Bond, R., & Smith, P. B. (1996). Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgment task. *Psychological Bulletin*, *119*(1), 111.

Bosworth, S., Kabay, M., & Whyne, E. (2014). *Computer security handbook* (6th ed.). New York: Wiley.

Bullée, J. H., Montoya, L., Junger, M., & Hartel, P. H. (2016). Telephone-based social engineering attacks: An experiment testing the success and time decay of an intervention. In Mathur, A., & Roychoudhury, A. (Eds.), *Proceedings of the Inaugural Singapore Cyber Security R&D Conference (SG-CRC 2016), Singapore, Singapore*, Cryptology and Information Security Series, Vol. *14*. Amsterdam: IOS Press, pp. 107–114.

Bullée, J. H., Montoya, L., Pieters, W., Junger, M., & Hartel, P. H. (2015). The persuasion and security awareness experiment: Reducing the success of social engineering attacks. *Journal of Experimental Criminology*, *11*(1), 97–115.

Chan, M., Woon, I., & Kankanhalli, A. (2005). Perceptions of information security at the workplace: Linking information security climate to compliant behavior.

Chikudate, N. (2009). If human errors are assumed as crimes in a safety culture: A lifeworld analysis of a rail crash. *Human Relations*, *62*(9), 1267–1287.

Chiu, Y-N, Leclerc, B., & Townsley, M. (2011). Crime script analysis of drug manufacturing in clandestine laboratories: Implications for prevention. *British Journal of Criminology*, *51*(2), 355–374.

Cialdini, R. (2009). *Influence*. New York:: HarperCollins.

Cialdini, R., Vincent, J. E., Lewis, S. K., Catalan, J., Wheeler, D., & Darby, B. L. (1975). Reciprocal concessions procedure for inducing compliance: The door-in-the-face technique. *Journal of Personality and Social Psychology*, *31*(2), 206–215.

Code of Federal Regulations (2005). Title 45: Public Welfare, Department of Health and Human Services, Part 46: Protection of Human Subjects.

Collins, N. L., & Miller, L. C. (1994). Self-disclosure and liking: A meta-analytic review.*Psychological Bulletin*, *116*(3), 457–475.

Cornish, D. (1994). The procedural analysis of offending and its relevance for situational prevention. *Crime Prevention Studies*, *3*, 151–196.

Dang, H. (2008). The Origins of Social Engineering. *McAffee Security Journal*, *1*(1), 4–8.

Dimkov, T., Pieters, W., & Hartel, P. H. (2009). Two methodologies for physical penetration testing using social engineering. (*No. TR-CTIT-09-48*). Enschede.

Dreyfus, S., & Assange, J. (2012). *Underground: Tales of hacking, madness and obsession on the electronic frontier*. Edinburgh: Canongate Books.

Edmondson, A. C. (1996). Learning from mistakes is easier said than done: Group and organizational influences on the detection and correction of human error. *The Journal of Applied Behavioral Science*, *32*(1), 5–28.

Feeley, T. H., Anker, A. E., & Aloe, A. M. (2012). The door-in-the-face persuasive message strategy: A meta-analysis of the first 35 years. *Communication Monographs*, *79*(3), 316–343.

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using r*. London: SAGE Publications.

Finn, P., & Jakobsson, M. (2007). Designing ethical phishing experiments. *IEEE Technology and Society Magazine*, *26*(1), 46–58.

Freedman, J. L., & Fraser, S. C. (1966). Compliance without pressure: The foot-in-the-door technique. *Journal of Personality and Social Psychology*, *4*(2), 195–202.

Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond heuristics and biases. *European Review of Social Psychology*, *2*(1), 83–115.

Greenspan, S. (2008). *Annals of gullibility: Why we get duped and how to avoid it*, Non-Series. Westport: Praeger.

Griskevicius, V., Goldstein, N. J., Mortensen, C. R., Sundie, J. M., Cialdini, R. B., & Kenrick, D. T. (2009). Fear and loving in Las Vegas: Evolution, emotion, and persuasion. *Journal of Marketing Research (JMR)*, *46*(3), 384–395.

Gupta, M., Agrawal, S. (2011). A Survey on Social Engineering and the Art of Deception. *International Journal of Innovations in Engineering and Technology*, *1*(1), 31–35.

Hadnagy, C., & Wilson, P. (2010). *Social engineering: The art of human hacking*. New York: Wiley.

Hofling, C., Brotzman, E., Dalrymple, S., Graves, N., & Pierce, C. (1966). An experimental study in nurse-physician relationships. *The Journal of nervous and mental disease*, *143*(2), 171.

Huber, M., Kowalski, S., Nohlberg, M., & Tjoa, S. (2009). Towards automating social engineering using social networking sites, *Computational Science and Engineering, 2009. CSE '09. International Conference on*, Vancouver, BC, Canada, Vol. *3*, pp. 117–124.

Janczewski, L., & Colarik, A. (2008). *Cyber warfare and cyber terrorism*, Gale virtual reference library: Information Science Reference, Hershey, PA.

Kennedy, D. (2011). There's something "human" to social engineering. http://magazine.thehackernews.com/article-.html.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1)159–174.

Lavorgna, A. (2014). Wildlife trafficking in the internet age. *Crime Science*, *3*(1), 5.

Luo, R. X., Brody, R., Seazzu, A., & Burd, S. (2011). Social engineering: The neglected human factor for information security management. *Information Resources Management Journal*, *24*(3), 1–8.

Mann, I. (2008). *Hacking the human: Social engineering techniques and security countermeasures*. Aldershot: Gower.

Marconato, G. V., Kaaniche, M., & Nicomette, V. (2012). A vulnerability life cycle-based security modeling and evaluation approach. *The Computer Journal*, *56*(4), 422–439.

Meyerowitz, B. E., & Chaiken, S. (1987). The effect of message framing on breast self-examination attitudes, intentions, and behavior. *Journal of Personality and Social Psychology*, *52*(3), 500–510.

Milgram, S. (1974). *Obedience to authority: An experimental view*. New York: Harper & Row.

Milgram, S. (1963). Behavioral study of obedience. *The Journal of Abnormal and Social Psychology*, *67*(4), 371–378.

Milgram, S. (1965). Some conditions of obedience and disobedience to authority. *Human Relations*, *18*(1), 57–76.

Mitnick, K., Simon, W. L., & Wozniak, S. (2011). *Ghost in the wires: My adventures as the world's most wanted hacker*. New York: Little, Brown.

Mitnick, K., & Simon, W. (2002). *The art of deception: Controlling the human element of security*. New York:Wiley.

Mouton, F., Malan, M. M., Kimppa, K. K., & Venter, H. (2015). Necessity for ethics in social engineering research. *Computers & Security*, *55*, 114–127.

Muscanell, N. L., Guadagno, R. E., & Murphy, S. (2014). Weapons of influence misused: A social influence analysis of why people fall prey to internet scams. *Social and Personality Psychology Compass*, *8*(7), 388–396.

O'Keefe, D. J., & Hale, S. L. (2001). An odds-ratio-based meta-analysis of research on the door-in-the-face influence strategy. *Communication Reports*, *14*(1), 31–38.

Pascual, A., & Guéguen, N. (2005). Foot-in-the-door and door-in-the-face: A comparative meta-analytic study 1. *Psychological Reports*, *96*(1), 122–128.

Poulsen, K. (2011). *Kingpin: How one hacker took over the billion-dollar Cybercrime underground*. New York: Crown/Archetype.

Reason, J. (1990). The contribution of latent human failures to the breakdown of complex systems. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *327*(1241), 475–484.

Rhee, H.-S., Kim, C., & Ryu, Y. U. (2009). Self-efficacy in information security: Its influence on end users' information security practice behavior. *Computers & Security*, *28*(8), 816–826.

Rouse, M. (2006). *Definition social engineering*: TechTarget. http://www.searchsecurity.techtarget.com/definition/social-engineering.

Rowe, E., Akman, T., Smith, R. G., & Tomison, A. M. (2012). Organised crime and public sector corruption: A crime scripts analysis of tactical displacement risks. *Trends and Issues in Crime and Criminal Justice*, *444*, 1.

Schellevis, J. (2011). Grote amerikaanse bedrijven vatbaar voor social engineering. http://tweakers.net/nieuws/77755/grote-amerikaanse-bedrijven-vatbaar-voor-social-engineering.html.

Schneier, B. (2000). *Secrets & lies: Digital security in a networked world* (1st ed.). New York, NY, USA: John Wiley & Sons, Inc.

Sutton, R. M., Niles, D., Meaney, P. A., Aplenc, R., French, B., Abella, B. S., ... Nadkarni, V. (2011). Low-dose, high-frequency cpr training improves skill retention of in-hospital pediatric providers. *Pediatrics*, *128*(1), e145–e151.

Szymanski, D. (2001). Modality and offering effects in sales presentations for a good versus a service. *Journal of the Academy of Marketing Science*, *29*(2), 179–189.

Tanford, S., & Penrod, S. (1984). Social influence model: A formal integration of research on majority and minority influence processes. *Psychological Bulletin*, *95*(2), 189–225.

The Federal Bureau of Investigation (2013). Internet Social Networking Risks (Vol. 2013)(No. 4 October). U.S. Department of Justice. Retrieved 23-Oktober-2013, from. http://www.fbi.gov/about-us/investigate/counterintelligence/internet-social-networking-risks.

Thompson, L., & Chainey, S. (2011). Profiling illegal waste activity: Using crime scripts as a data collection and analytical strategy. *European Journal on Criminal Policy and Research*, *17*(3), 179–201.

Tremblay, P., Talon, B., & Hurley, D. (2001). Body switching and related adaptations in the resale of stolen vehicles. Script elaborations and aggregate crime learning curves. *British Journal of Criminology*, *41*(4), 561–579.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131.

Twitchell, D. P. (2009). Social engineering and its countermeasures, *Handbook of Research on Social and Organizational Liabilities in Information Security*: Hershey, PA: IGI-Global, pp. 228–242.

USA v. Mitnick (1996). Indictment, CR 96-881, 145 F.3d 1342.

USA v. Mitnick (1998). No. 97-50365.

Whittingham, R. (2004). *The blame machine: Why human error causes accidents*. London: Taylor & Francis.

Winkler, I. S., & Dealy, B. (1995). Information security technology? ... don't rely on it: A case study in social engineering, *Proceedings of the 5th Conference on Usenix Unix Security Symposium - Volume 5*. Berkeley, CA, USA: USENIX Association, pp. 1–1.

Zhao, B., & Olivera, F. (2006). Error reporting in organizations. *Academy of Management Review*, *31*(4), 1012–1030.

How to cite this article: Bullée J-WH, Montoya L, Pieters W, Junger M, Hartel P. On the anatomy of social engineering attacks—A literature-based dissection of successful attacks. *J Investig Psychol Offender Profil*. 2018;15:20–45. https://doi.org/10.1002/jip.1482