J Exp Criminol (2015) 11:97–115
DOI 10.1007/s11292-014-9222-7

The persuasion and security awareness experiment: reducing the success of social engineering attacks

Jan-Willem H. Bullée · Lorena Montoya · Wolter Pieters · Marianne Junger · Pieter H. Hartel

Published online: 20 January 2015

Abstract *Objectives* The aim of the current study is to explore to what extent an intervention reduces the effects of social engineering (e.g., the obtaining of access via persuasion) in an office environment. In particular, we study the effect of authority during a 'social engineering' attack. *Methods* Thirty-one different '*offenders*' visited the offices of 118 employees and on the basis of a script, asked them to hand over their office keys. Authority, one of the six principles of persuasion, was used by half of the offenders to persuade a target to comply with his/her request. Prior to the visit, an intervention was randomly administered to half of the targets to increase their resilience against attempts by others to obtain their credentials. *Results* A total of 37.0 % of the employees who were exposed to the intervention surrendered their keys while 62.5 % of those who were not exposed to it handed them over. The intervention has a significant effect on compliance but the same was not the case for authority. *Conclusions* Awareness-raising about the dangers, characteristics, and countermeasures associated with social engineering proved to have a significant positive effect on neutralizing the attacker.

Keywords Authority · Awareness · Credentials · Experiment · Intervention · Persuasion · Social engineering

J. H. Bullée (✉) · L. Montoya · W. Pieters · P. H. Hartel
Services, Cyber-security, and Safety Group (SCS), Faculty of EEMCS, University of Twente, PO Box 217, Enschede 7500 AE, The Netherlands
e-mail: j.h.bullee@utwente.nl

M. Junger
Industrial Engineering and Business Information Systems (IEBIS), Faculty of Management and Governance, University of Twente, PO Box 217, Enschede 7500 AE, The Netherlands

W. Pieters
ICT section, Faculty of Technology, Policy and Management, Delft University of Technology, PO Box 5015, Delft, 2600 GA, The Netherlands

Introduction

Many attempts at unauthorized access involve exploitation of human weaknesses. Whereas this so-called social engineering had previously been associated with the social sciences, its terminology has caught on among computer and information security professionals (Anderson 2008). Social engineering is a non-technical type of attack based on human interaction and often involves tricking others into breaking security policies. Examples of social engineering include persuading targets to run malware-infected e-mail attachments as well as phishers trying to convince people to disclose sensitive information. In the field of phishing, e-mails often lead to fake websites that resemble legitimate websites, and lure people into disclosing sensitive information (Kumaraguru et al. 2010). Security experts propose that as our culture becomes more dependent on information technology, social engineering will become the greatest threat to any security system (Rouse 2006).

These examples show that social engineers know that humans are a weak link in cyber security, and therefore try to trick people into violating security policies. The actions of social engineers are designed to appear harmless and to look legitimate (The Federal Bureau of Investigation 2013). For some time, social engineering has been part of many organizational security testing programs. Such programs typically involve an outside party conducting (a) technical tests such as vulnerability assessments, penetration tests, audits, and (b) user-oriented tests such as social engineering and phishing tests. These actions can be simple ones such as calls using invented scenarios (i.e., pretexting, tailgating) or complex such as sophisticated make-up and appropriate acting. Many organizations choose to conduct both on-site and remote social engineering testing to elevate the staff's ability to recognize social engineering attacks and to find out how they respond (Cross 2011). However, little scientific literature has shown how persuasion techniques influence the success of a social engineering attack or the effect of countermeasures. This paper therefore explores to what extent (a) persuasion techniques influence the outcome and (b) an intervention reduces the effects of social engineering.

Persuasion

Humans are susceptible to persuasion by nature and in some circumstances resisting is almost impossible. The theory of gullibility explains the susceptibility to persuasion and it is an extension of credulity, i.e., the willingness to believe someone or something in the absence of reasonable proof (Greenspan 2008). Gullibility adds a concrete action to credulity and it results in a negative outcome. A gullible person believes in the goodness of people with the unfortunately negative consequence of being duped even in the face of warning signs (Greenspan 2008). The theory of optimistic bias states that people believe that positive events are more likely to occur to them than to other people (Weinstein 1980). The inverse is also true: people believe that negative events are more likely to occur to other people than to themselves. In general, even when multiple attempts are made to reduce it, optimistic bias remains (Harris et al. 2000). Optimistic bias together with gullibility implies that people think

that they (a) will not be selected as a social engineering target and (b) are more likely to resist than others.

Once a person is a target, the offender can use persuasion techniques to change the odds in his/her favor. Six principles of persuasion can be used to increase the offender's probability of success: reciprocity, conformity, liking, scarcity, commitment and authority (Cialdini 2009). *Reciprocity* refers to the giving of something in return. The target feels indebted to the requester for making a gesture. Even the smallest gift puts the requester in an advantageous position. *Conformity*, or social proof, is imitating the behavior of other people. Members of the in-group have a stronger feeling of group-safety compared with members of the out-group (Asch 1951). *Liking* someone puts that person in a favorable 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. *Commitment* refers to the likelihood of sticking to a cause or idea after making a promise or agreement. In general, when a promise is made, people will honor it, which increases the likelihood of compliance (Cialdini 2009).

Authority is the principle that describes people's tendency to obey the request of authoritative figures. If people are unable to make a well-informed decision, the responsibility to do so is transferred to the group or person they believe is in charge. Crisis and stress activate the behavioral trait of responsibility transition. The Authority principle can be operationalized in multiple ways; the three most common are: (1) prestigious titles such as professor, doctor, or lawyer give strength to an argument of authority; (2) stylish and expensive outfits carry an aura of status and position, as do trappings such as jewelry and cars; (3) the physical appearance of a person is an indicator of authority (Doob and Gross 1968). There are differences in perception of authority towards someone wearing a police uniform, a casual outfit, a business suit or work clothing (Bickman 1974; Lefkowitz et al. 1955).

The most famous study illustrating authority is the classical experiment of Stanley Milgram, which tested obedience to authority. When they were instructed to do so by a man wearing a lab coat, 66 % of the participants did not hesitate to administer a dose of 450 V to a human test subject (Milgram 1963). Replication studies and meta-analysis found comparable results (Burger 2009; Packer 2008). An evaluation of 23 replication studies, conducted during the past 35 years, studied the obedience to authority paradigm. Almost half of the studies (11) showed a lower rate of obedience ranging between 28 % and 65 %. The remaining 12 studies showed an equal or higher rate of obedience (Blass 1999). Although the studies found different rates of obedience, it can be concluded that authority is an important behavioral phenomenon.

The operationalization of authority using clothing was demonstrated by means of an experiment (Bickman 1974; Lefkowitz et al. 1955). A stranger requested a small donation for the payment of a parking meter. When a police officer asked for a contribution, 92 % were willing to contribute, compared to 42 % in the case of a civilian (Bickman 1974). A difference also was found between the number of people willing to follow someone who was crossing illegally at an intersection, depending on

whether the person was casually or formally dressed. If the pedestrian was wearing a well-tailored suit, 3.5 times more people followed and crossed illegally than when the pedestrian wore casual clothing (Lefkowitz et al. 1955). The authority principle, operationalized by formal clothing, will be used as experimental condition in this study.

Behavioral changes to countermeasure persuasion

As stated before, offenders are aware of the susceptibility of humans to persuasion. Interventions are a common way to promote behavioral change and make potential targets more resistant against psychological manipulation. A wide range of approaches, strategies, and theories can be used. Many of them are well studied and are the result of extensive research. Well-known strategies and theories on attitude change include: (1) shift of focus from the Protection Motivation Theory (Rogers 1975), (2) the effect of social comparison shown by Festinger (1957), (3) correction of misconception from the Theory of Planned Behavior (Ajzen 1988, 1991), (4) model learning by observing role models from the Social Learning Theory (Bandura 1986), (5) persuasive communication from the Elaboration Likelihood Model (ELM) (Petty and Cacioppo 1986). The ELM of Persuasion describes how information is processed and can be tailored to the receivers. According to the ELM, people process information via either the peripheral or the central route. The peripheral route of information processing is used when there is minimal attention to the message and can involve superficial cues, such as the attractiveness of the message presented. One may like the sound of a person's voice, or that person might have gone to the same university as one did. The central route, on the other hand, involves persuasion on the basis of the message content, such as voting for the political party with the best arguments (Petty and Cacioppo 1981). Consistent with the ELM theory, attitudes obtained via the central route last longer, are less vulnerable to contra-argumentation and are better predictors of human behavior. Furthermore, the effect of persuasive communication increases if the message is relevant to the audience and if surprises and repetitions are used (Petty and Cacioppo 1984, 1986).

Persuasive communication by means of leaflets has been thoroughly studied and proved to be an effective mechanism for administering an intervention. Studies about successful leaflets have focused on increasing the knowledge of the general public (Humphris et al. 1999; Stubbings et al. 2000), as well as more specific groups such as patients (Barlow 1998; Hawkey and Hawkey 1989), parents (Ghaderi et al. 2013), and customers (Shim et al. 2011). Other studies found that leaflets influenced behavior (Ershoff et al. 1989; Hart et al. 1997) and reduced anxiety about an illness (Humphris et al. 2001; Robb et al. 2006). However, conflicting results were reported by Carré et al. (2008). They argued that the readability of their leaflet was probably not optimal for the audience, although the majority of the subjects found the leaflet fairly clear and interesting (Carré et al. 2008).

A meta-analysis on smoking cessation compared the distribution of leaflets against a control condition (i.e., no leaflet). In total, 12 trials were conducted and a significant effect was found on smoking cessation ($N = 14{,}787$; $p = .008$) (Lancaster and

Stead 2005). Similarly, experts argue that both training and education can help protect users against phishing (Hight 2005; Kumaraguru et al. 2010). Intervention materials currently available can be effective as long as the user actually reads the material (Kumaraguru et al. 2010). In the field of advertising, a distinction can be made between comparative and non-comparative advertisements; the latter are peripherally processed (Lien 2001). A meta-analysis (d = effect size) on advertisements showed that comparative ads led to an increase in brand attitude ($N = 42$; $d = .23$), purchase intention ($N = 47$; $d = .20$), and purchase behavior ($N = 6$; $d = .46$) (Grewal and Kavanoor 1997). Information leaflets may offer an important contribution to raising long-term knowledge and awareness, but to be effective, these must be well presented and understandable by the target population (Krawczyk et al. 2012; Petty and Cacioppo 1986).

Subtle reminders have shown to be a useful mechanism for remembering the contents of a leaflet within the context of promoting desired behavior (Gisquet-Verrier and Riccio 2012; Glanz et al. 1997). Reminders (cues to action) have proved to be necessary in the promotion of healthy behavior (Rosenstock 1974). Reminders can either be internal or external (e.g., medical advice, postcard reminder, television advertisement, warning label). Reminders are mentioned in the theory of Situational Crime Prevention as elements that Remove Excuses (Cornish and Clarke 2003). Although the literature on reminder cues is limited, research suggests that these are effective (Flight et al. 2012).

Humor has been found to have a positive effect on the recall of information compared to neutral cues (Carlson 2011, p. 75; Gulas and Weinberger 2006). Schmidt (1994) describes three mechanisms that explain the effect of humor on information recollection. The first considers the effect of humor on physiological arousal. Humor is associated with an increase in heart and respiratory rates (Schmidt 1994). Increased arousal rates during the presentation of humorous material are found to lead to long-term memory retention, compared to neutral arousal rates (Craik and Blankstein 1975). Second, it is argued that humor increases attention towards a subject. Paying more attention to something therefore results in better recollection (Schmidt 1994). A third mechanism is the effect of repetition caused by humorous material. Recollection is better when material is presented on more than one occasion (Schmidt 1994).

Experimental context

Literature shows that social engineering works (Mann 2008), that it is effective (Schellevis 2011), and that targets are unaware of being victimized (The Federal Bureau of Investigation 2013; Hadnagy and Wilson 2010). However, there is limited literature on success rates for such attacks. The books of Kevin Mitnick (2002; 2011) give an insight into how successful social engineering attacks are executed. Although this is interesting anecdotal information, the prevalence and success rate of social engineering attacks are unknown. Penetration testing reports occasionally surface, but these represent the proverbial needle in a haystack and are uncontrolled. To the best of our knowledge, there are no other studies that combine persuasion principles with an intervention. Therefore, an experiment was conducted to measure the

success of social engineering attacks in a controlled setting. Only the success rate was measured, prevalence was not taken into account.

The success of a social engineering attack might be substantially influenced by the context. Context can influence and can be influenced by both the target and the offender to obtain a more desirable outcome. The offender may use the knowledge of the six principles of persuasion to achieve a higher probability of compliance. On the other hand, the target can be informed about social engineering via an intervention aiming at rejecting the offender's requests. In the current experiment, the context influenced by the offender is authority, whilst the context influenced by the defender is the intervention. The social engineering experiment was administered to university personnel with the goal of making them surrender their office key. The advantage of a key over a password is its appearance, since a key is physical and a password represents knowledge. It is therefore possible to surrender a key, but a password has to be shared. One cannot return a password without still knowing it, this additionally introduces the burden of changing one's password afterwards.

Research question

The objective of this research was to answer the following question: "*To what extent are people susceptible to social engineering attacks*?" Three hypotheses were formulated: *H1*) Previous research on the effect of informing people showed an increase of their knowledge and a change in behavior. We therefore hypothesize that in the intervention group, fewer people comply with the offender's request than in the control group. *H2*) Previous research on the effect of authority showed an increase in compliance towards an authoritative figure compared to non-authoritative figures. We therefore hypothesize that in the experimental group, where the researcher exercises authority (i.e., by wearing formal clothing), more people comply with the offender's request than in the control group. *H3*) An intervention reduces the effect of authority in relation to the compliance with a request and therefore the third hypothesis is that authority has more effect if no intervention is exercised.

Method

The sample consisted of 118 subjects of both sexes who work in two buildings on the University of Twente's campus. Only people who were present in their office and whose office door had a specific type of lock were approached. The locks are electronic and manufactured by WinkHaus. A physical plastic token (resembling a key) with a chip inside is used to unlock the door (refer to Fig. 1). The chip contains data that specifies which locks are allowed to be opened. A 128-bit challenge-response between the key (chip) and the lock is used to validate the identity of the key. When the key is authorized, the lock pulls a small pin and allows one to open the door (Schneier 2005). The validity of the credentials on the key lasts until midnight and thus the key has to be reactivated on a daily basis. The activation facilities are located next to the entrance to the buildings so that staff can activate the key upon arrival. Staff who forgets to activate must return to the entrance.

Fig. 1 Key

Subjects

Professors, secretaries, and laboratory staff were excluded from the experiment in order to minimize disruption of the main university activities. The sample only consisted of one-third of all possible targets; the gender distribution is comparable to the overall gender distribution of the buildings, while the experimental sample is slightly younger (34 vs. 40 years). Although there are some small differences with respect to a number of characteristics, the sample deviates little from the general population of these buildings.

Researchers

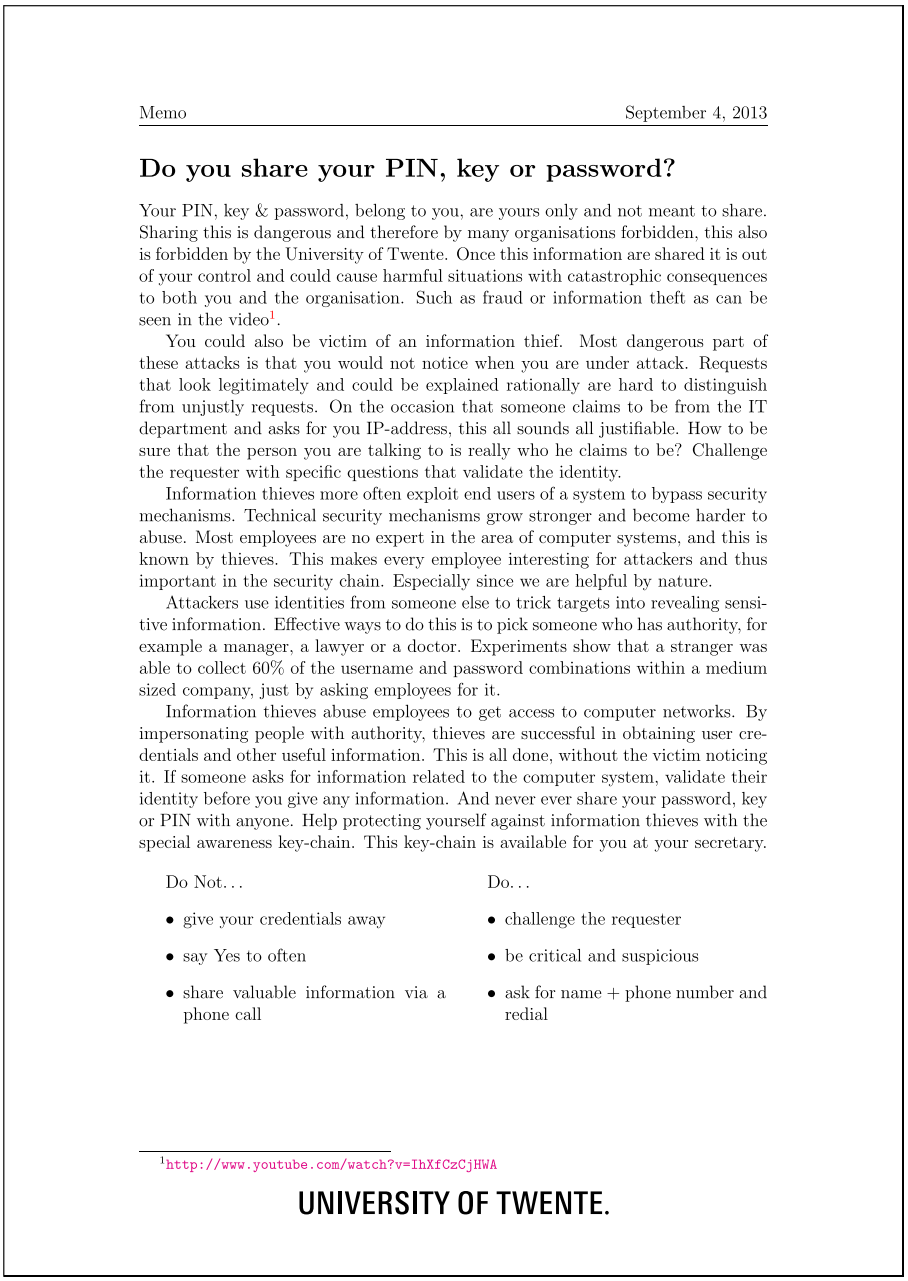
The researchers (i.e., the "attackers") consisted of 31 bachelor/master students (9 female and 22 male). The average age of the researchers was 21.26 years (SD = 1.34). There was no restriction in terms of approaching subjects of the same gender.

Procedure

One week before the researchers approached the subjects, half of the subjects were exposed to an intervention. Departments were randomly selected and all their staff was exposed to the intervention, in an effort to avoid disturbing the normal activities of the department.

No targets from the control group reported having been aware that an intervention had taken place, whereas those in the intervention group reported to have received the intervention. This was ascertained during the debriefing when subjects were directly asked. The target group consisted of university scientific staff and it was assumed that all of them were capable of processing the arguments presented in the intervention.

The intervention consisted of: (a) a leaflet explaining what social engineering is, why is it dangerous, how to detect it and what to do if confronted with it (refer to Fig. 2), (b) a blue key chain with the university logo on one side and the text "Don't give me to a stranger" on the other side (refer to Fig. 3a and b respectively) and (c) a poster containing a humorous quote and an explicit remark against password, key and PIN sharing (refer to Fig. 4). The leaflets were designed



Memo September 4, 2013

Do you share your PIN, key or password?

Your PIN, key & password, belong to you, are yours only and not meant to share. Sharing this is dangerous and therefore by many organisations forbidden, this also is forbidden by the University of Twente. Once this information are shared it is out of your control and could cause harmful situations with catastrophic consequences to both you and the organisation. Such as fraud or information theft as can be seen in the video[1].

You could also be victim of an information thief. Most dangerous part of these attacks is that you would not notice when you are under attack. Requests that look legitimately and could be explained rationally are hard to distinguish from unjustly requests. On the occasion that someone claims to be from the IT department and asks for you IP-address, this all sounds all justifiable. How to be sure that the person you are talking to is really who he claims to be? Challenge the requester with specific questions that validate the identity.

Information thieves more often exploit end users of a system to bypass security mechanisms. Technical security mechanisms grow stronger and become harder to abuse. Most employees are no expert in the area of computer systems, and this is known by thieves. This makes every employee interesting for attackers and thus important in the security chain. Especially since we are helpful by nature.

Attackers use identities from someone else to trick targets into revealing sensitive information. Effective ways to do this is to pick someone who has authority, for example a manager, a lawyer or a doctor. Experiments show that a stranger was able to collect 60% of the username and password combinations within a medium sized company, just by asking employees for it.

Information thieves abuse employees to get access to computer networks. By impersonating people with authority, thieves are successful in obtaining user credentials and other useful information. This is all done, without the victim noticing it. If someone asks for information related to the computer system, validate their identity before you give any information. And never ever share your password, key or PIN with anyone. Help protecting yourself against information thieves with the special awareness key-chain. This key-chain is available for you at your secretary.

Do Not...

- give your credentials away
- say Yes to often
- share valuable information via a phone call

Do...

- challenge the requester
- be critical and suspicious
- ask for name + phone number and redial

[1] http://www.youtube.com/watch?v=IhXfCzCjHWA

UNIVERSITY OF TWENTE.

Fig. 2 Memo used in the intervention

so that they could be processed via the central route, which was achieved by the absence of peripheral cues in it (Petty and Cacioppo 1986). The leaflet represents the

(a) Front View (b) Back View

Fig. 3 Key chain

information media, the key chain the subtle reminder of the intervention and the humorous poster the cue that helps to remember the leaflet better. Departmental secretaries were responsible for distributing the material; they were unaware that this was part of an experiment. They were only instructed to distribute the material within their research group.

The leaflet and the poster were distributed by e-mail, while the key chain was distributed in person. It is unknown whether any of the subjects printed the posters and displayed them in their office; however, there were no reports of the intervention material being displayed in public areas (e.g., hallways, coffee corners, or lunchrooms). All subjects were individually approached by a researcher between 10 a.m. and 6 p.m. on a 'normal' Wednesday during term time. In order to avoid suspicion, researchers (i.e., attackers) never made consecutive visits to members of the same department. After each visit, they therefore had to come back to the base of operations (i.e., the first author's office) to obtain the name and location of the next target, which was randomly selected from a list of all possible targets.

The researchers were randomly assigned to either the authority or the control condition, that is, wearing formal or casual clothing, respectively. Subjects were randomly assigned to one of four conditions: they were either exposed to (a) both authority and intervention, (b) authority but not intervention, (c) intervention but not authority, or (d) they were exposed to neither authority nor intervention.

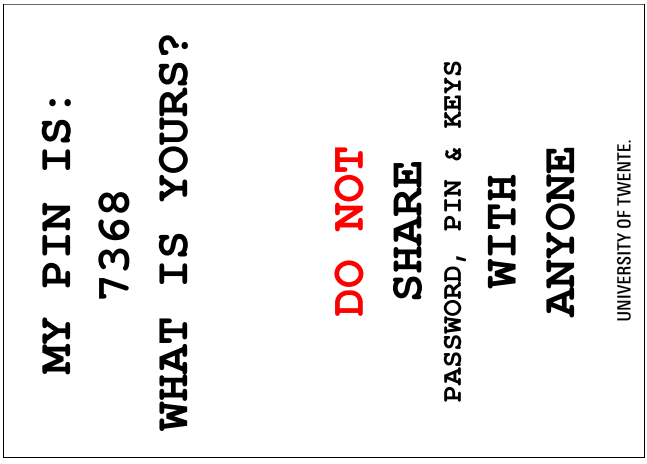


Fig. 4 Poster used in the intervention

Each researcher approached the subjects in their office using the following script:

> Hi, I am `[Name]` and I work for Facility Management. I have a question regarding the door locks. We received several complaints about the door lock and the keys. Has unlocking the door ever been problematic for you? We have contacted the manufacturer about the malfunctioning and they had received other similar complaints. In order to solve the problem, the manufacturer sent us a measuring device to test the keys that are in use. I have to admit that I don't exactly know what the box measures, but the data collected is necessary for the manufacturer to analyze the situation and hopefully find a solution to the problem. Can I have your key for measurement?
> *After measuring the key*: I have to inform you that after reading your key, the key has been reset and needs reactivation downstairs. It is no problem for me to reactivate the key for you.
> **Request**: Is it OK with you if I do the reactivation of your key downstairs?

Each target was subjected to the same request. After the researcher obtained the key and walked away, he/she came back to return the key and orally debriefed the subject with regards to social engineering. During the debriefing session, the subject was asked some demographic information, their opinion of physical and digital security at the university, length of employment, their route to the activation point and, for those who had not complied, to explain why they had not handed the key over.

Variables

The variables used in the analysis were: compliance, intervention, and authority. The dependent variable compliance measured whether the subject complied with the request of the offender to hand over the key. The dichotomous variable was dummy coded as 0 = did not comply, 1 = did comply. The independent variable intervention measured whether the subject was exposed to the intervention (0 = not exposed to the intervention, 1 = exposed to the intervention). The independent variable authority measured whether the offender (i.e., research assistant) wore casual clothing or wore formal clothing. Casual clothing was operationalized by wearing jeans and a t-shirt and formal clothing by a buttoned collar shirt and trousers (coded as 0 = informally dressed, 1 = formally dressed).

Analysis

The first two hypotheses were tested using cross tabulations and a Chi-square test. For the third hypothesis, logistic regression was used. The following three assumptions must be met for logistic regression analysis: (1) sufficient sample size, (2) uncorrelated independent variables, and (3) outlier-free dataset (Pallant 2010). The VIF (Variance Inflation Factor) is 1.002 for both the authority and intervention variables, which is below the cut-off value of 10, indicating that there is no evidence of multicollinearity (Pallant 2010). Only dichotomous variables were used, meaning that all variables are either 0 or 1, thus there are no outliers. To find out if the intervention

Table 1 Number of observations and percentages per intervention condition

		Intervention		
		No	Yes	Total
Complied	No	27 (37.5 %)	29 (63.0 %)	56 (47.5 %)
	Yes	45 (62.5 %)	17 (37.0 %)	62 (52.5 %)
Total		72 (100 %)	46 (100 %)	118 (100 %)

affects the relation between authority and compliance, a moderation analysis was carried out.

Compliance was predicted on the basis of (a) model 1: (I) for intervention, (b) model 2: (A) for authority and (c) model 3: (AxI) for authority, intervention, and the interaction between the two. The third model is needed to test whether the simultaneous influence of two variables on a third is non-additive. The relevance is that if authority and intervention interact, the relationship between each of the interacting variables and compliance depends on the value of the other interacting variable.

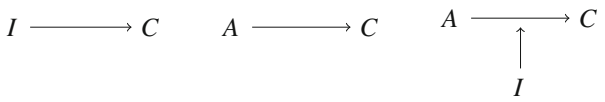
Finally, the three models (refer to Fig. 5a, b and c) were compared to each other on the basis of the variance and likelihood ratio.

Results

A total of 118 subjects were approached. No 'building occupied by target' effect on compliance ($N = 118$, $df = 1$, $\chi^2 = .514$, $p = .473$), 'offender's gender' effect on compliance ($N = 118$, $df = 1$, $\chi^2 = .961$, $p = .327$) or 'target's gender' effect on compliance ($N = 118$, $df = 1$, $\chi^2 = .279$, $p = .586$) were found and these are therefore not further mentioned.

Hypothesis 1: Subjects in the intervention and control groups comply unequally

Of the targets who were not exposed to the intervention (control group), 62.5 % agreed to the request of the offender compared to 37 % of those in the intervention group ($\chi^2 = 7.34$, $df = 1$, $p = .007$). Hypothesis *H1* is therefore accepted. Those in the control group have 2.84 times higher odds of compliance (i.e., handing over their keys) than those who were exposed to the intervention ($OR = 2.84$, CI [1.32, 6.1]). Refer to Table 1 and Fig. 6a for descriptive statistics.



(a) Model 1 (b) Model 2 (c) Model 3

Fig. 5 Overview of the three models

Hypothesis 2: Subjects in the authority and control groups comply unequally

In total, 53.6 % of the targets in control group (those approached by casually dressed offenders) complied compared to 51.6 % of those in the authority group ($\chi^2 = 0.45$, $df = 1$, $p = .832$). *H2* is rejected in favor of the alternative hypothesis *H2a*: "Authority and Control comply equally". Those in the control group have 1.08 times higher odds of compliance (i.e., handing over their keys) than those that were exposed to authority ($OR = 1.08$, CI [0.52, 2.23]). Refer to Table 2 and Fig. 6b for descriptive statistics.

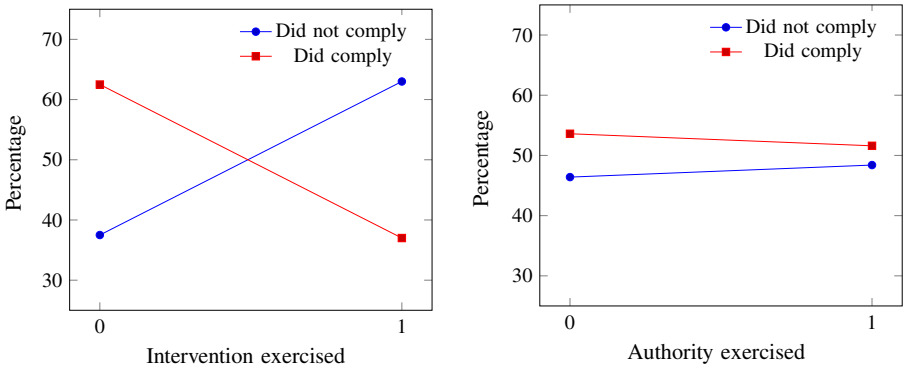
Hypothesis 3: An intervention reduces the effect of authority in relation to the compliance with a request

It was tested whether the relation between authority and compliance was affected by the intervention. In the model with the interaction term (refer to Model 3: AxI in Table 3 and Fig. 7), this variable is marginally significant ($p = .093$). The hypothesis *H3* is rejected in favor of the alternative hypothesis *H3a*, thus there is only a tendency for the intervention to moderate (i.e., affect) the relation between authority and compliance. Refer to Table 3 and Fig. 7 for multiple regression results.

The comparison of the three models (refer to Table 3) shows a significant difference between model 1 & 2 and between 2 & 3. The pseudo R^2 shows zero percentage of explained variance for model 2 (authority), while in model 1 (intervention) this number increased.

Discussion

This study investigated whether an intervention in the form of an awareness campaign affects the relationship between the authority principle of persuasion and compliance



(a) Predicted change of compliance due to the intervention.

(b) Predicted change of compliance due to authority.

Fig. 6 Predicted change in compliance

Table 2 Number of observations and percentages per authority condition

		Authority		
		No	Yes	Total
Complied	No	26 (46.4 %)	30 (48.4 %)	56 (47.5 %)
	Yes	30 (53.6 %)	32 (51.6 %)	62 (52.5 %)
Total		56 (100 %)	62 (100 %)	118 (100 %)

with the request to hand over the office key to a stranger impersonating a facility management staff member.

An intervention composed of (a) informing people about the risks of social engineering attacks, (b) distributing a small key chain, and (c) a humorous poster, proved to have a large impact on the likelihood of handing over office keys to strangers.

In total, 37 % of staff exposed to the intervention versus 62.5 % in the control group complied with the request to hand over their keys. Not exposing staff to the intervention means that the odds of them handing over their keys are 2.84 times higher. The priming effect is in line with the results from studies in the medical sciences (Lancaster and Stead 2005) and studies on Internet-related crimes (Ferguson 2005; Hight 2005; Kumaraguru et al. 2010). The effect of leaflets has proven to be effective in the fields of promotion of healthy behavior, advertisement, and phishing (Lancaster and Stead 2005; Grewal and Kavanoor 1997; Kumaraguru et al. 2010). Existing literature shows that the processing of information in leaflets tends to cause behavioral change (e.g., healthier behavior, increased purchase actions, or a lower response rate to phishing mails). Our research confirms that informing people also leads to a change in behavior. Knowledge about social engineering

Table 3 Model comparison of the three models. The columns depict for each variable: the odds ratio (OR), its lower and upper 95 % confidence intervals [*in brackets*] and its significance level

	Model 1: (I)	Model 2: (A)	Model 3: (AxI)
Intervention	0.35 [0.16, 0.76]**		0.33 [0.15, 0.74]**
Authority		0.93 [0.45, 1.89]	0.73 [0.33, 1.61]
Auth*Int			0.26 [0.05, 1.25][a]
Constant	1.04 [0.65, 1.41]	1.11 [0.77, 1.59]	0.96 [0.63, 1.39]

* $p < .05$; ** $p < .01$; *** $p < .001$;
[a] $p = .093$;
Model 1 ($\chi^2 = 0.05$, $p < .007$), $N = 118$, pseudo $R^2 = .045$;
Model 2 ($\chi^2 = 7.41$, $p < .832$), $N = 118$, pseudo $R^2 = .000$;
Model 3 ($\chi^2 = 10.40$, $p < .016$), $N = 118$, pseudo $R^2 = .064$;
Model 1=2 ($p < .001$); Model 1=3 ($p < .224$); Model 2=3 ($p < .006$);

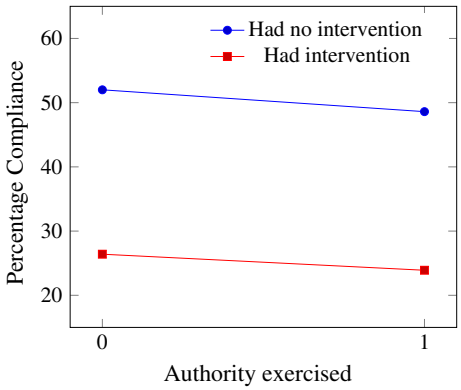


Fig. 7 The effect of intervention on authority

attacks helps to increase the safety of the organization and decreases susceptibility to victimization.

Allowing offenders to make changes to the context by applying principles of persuasion should increase their probabilities of success. However, in our study the expected effect of authority, operationalized via formal clothing, was not validated. Facility personnel at this particular university are formally dressed (since they wear black trousers and a blue shirt). It was therefore expected that informally dressed attackers would be viewed as having less power.

A possible explanation could be that countries differ in the perception of authority. To the best of our knowledge, the related research has all been carried out in the United States (Cialdini 2009; Bickman 1974; Lefkowitz et al. 1955; Milgram 1963). It is possible that Dutch people are less sensitive towards authority than Americans. However, there is no evidence of such variation, with respect to sensitivity to authority. Research on cultural differences shows that both countries score equally (Hofstede et al. 2010).

Hofstede identified six cultural dimensions and made a comparison across 60 countries. The six dimensions are: (a) indulgence versus restraint, (b) power distance, (c) individualism versus collectivism, (d) uncertainty avoidance, (e) long-term versus short-term orientation, and (f) masculinity versus femininity (Hofstede et al. 2010). The cultural dimension 'power distance' was thought to be one that could explain the perception of hierarchical structures and, accordingly the difference between the former and present findings. Three aspects indicate obedience to authoritative figures as in (Bickman 1974; Lefkowitz et al. 1955; Milgram 1963): (1) "Hierarchy in organizations reflects the existential inequality between higher-ups and lower-downs", (2) "Subordinates expect to be told what to do", and (3) "Parents teach children obedience" (Hofstede et al. 2010). Citizens from the United States and the Netherlands score similarly on power distance (Hofstede et al. 2010) and consequently, differences between countries cannot explain the lack of effect of

authority in the present study. However, in the present study, not all staff was Dutch. Future research should involve controlling for the country of origin of subjects.

Perhaps the experimental setting of the present study (i.e., a university) explains the lack of effect of authority. In addition, this university has a hierarchical structure that is somewhat 'flat'; the distance between the 'ranks' is relatively small. The contact between staff and their supervisors is on a first name basis and it is common practice to walk into someone's office without an appointment. In contrast, studies from the United States found in the literature took place in real-life situations on the street (Bickman 1974; Lefkowitz et al. 1955) or in a publicly available laboratory (Milgram 1963).

A second alternative explanation might relate to the experimental design. Subjects who volunteer to participate in experiments might experience authority differently from a psychological point of view compared to those who are forced to participate. Volunteer subjects might have developed a sense of commitment towards those running the experiment (e.g. the researcher). According to (Milgram 1974), commitment is the force that binds both the subject and authority to their role. The design of our experiment was different to the electro shock experiment, in which all subjects volunteered to participate (Milgram 1963). The same reasoning applies to the illegal intersection crossing experiment (Lefkowitz et al. 1955), where the subjects volunteered to participate. In our experiment, on the other hand, subjects were randomly selected and 'authority' was applied directly upon entrance into the office. The subjects therefore had less psychological binding and commitment to the authority, and thus were less likely to comply with the request.

When authority and intervention were entered together in a model, authority still lacked predictive power. It is possible that this was caused by the age of the offender. The average age of the researchers was slightly above 20 years old, whilst studies from previous research indicated that the experimental authoritative figure was over 31 years old (Bickman 1974; Lefkowitz et al. 1955; Milgram 1963).

Given that 62.5 % of staff handed over their office key, it seems that the currently used policies on this topic are in need of either clarification or proper dissemination. The results of the experiment were discussed with security and facility management personnel, who were surprised about the compliance rates. Moreover, the assumption that there is no need to make a policy explicit, since it is common sense not to give away bicycle, home or car keys to strangers, was found to be wrong.

Finally, we present recommendations for future research. First, the intervention encompasses three components: (a) a leaflet informing people about the risks of social engineering attacks; (b) a small key chain and (c) a humorous poster. Because each target in the intervention group was exposed to all three components, the measurement of the individual effect sizes was outside the scope of the current research and is hence considered as future research. The effect of each individual component (i.e. the key chain, the leaflet, the humorous poster) and their combination, should be tested in order to maximize the effectiveness of social engineering interventions.

Second, the decision to hand over one's office key, under the presumption that it is currently deactivated, depends on many contextual factors. In many cases, maintaining security implies additional costs in terms of time, money or effort. The decision to comply with the request of an offender could depend on the costs or efforts to be invested by oneself. A research question could be: 'how does the perception of safety and willingness to be safe change if the invested effort increases?' In practice this could be the time and effort it takes to reactivate the office key yourself.

Furthermore, a follow-up study in the form of a time-series analysis could be carried out to test whether the intervention effect is maintained over a period of time. If the effect does not hold, an evaluation of the time decay could provide practitioners in the field of security with relevant information for scheduling interventions. Furthermore, there is also a need to evaluate the effectiveness of reminders over time.

Finally, it is possible that the likelihood of handing over the office key is related to the level of security inside the office. An extension of the study could involve identifying which of the staff have security features installed such as Kensington locks or locked cabinets.

Acknowledgments 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.

References

Ajzen, I. (1988). *Attitudes, personality, and behavior (Mapping social psychology series)*. Dorsey Press.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. doi:10.1016/0749-5978(91)90020-T.

Anderson, R. (2008). *Security engineering: a guide to building dependable distributed systems*. Wiley.

Asch, S.E. (1951). Effects of group pressure upon the modification and distortion of judgments. In H. Guetzkow (Ed.), *Groups, Leadership, and Men* (pp. 177–190). Pittsburgh, PA: Carnegie Press.

Bandura, A. (1986). *Social foundations of thought and action (First Printing)*. Prentice Hall.

Barlow, J. (1998). Knowledge in patients with rheumatoid arthritis: a longer-term follow-up of a randomized controlled study of patient education leaflets. *Rheumatology*, *37*(4), 373–376. doi:10.1093/rheumatology/37.4.373.

Bickman, L. (1974). The social power of a uniform1. *Journal of Applied Social Psychology*, *4*(1), 47–61. doi:10.1111/j.1559-1816.1974.tb02599.x.

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. doi:10.1111/j.1559-1816.1999.tb00134.x.

Burger, J.M. (2009). Replicating Milgram: would people still obey today? *The American Psychologist*, *64*, 1–11. doi:10.1037/a0010932.

Carlson, K.A. (2011). The impact of humor on memory: is the humor effect about humor? *Humor - International Journal of Humor Research*, *24*(1). doi:10.1515/humr.2011.002.

Carré, P.C., Roche, N., Neukirch, F., Radeau, T., Perez, T., Terrioux, P., Ostinelli, J., Pouchain, D., Huchon, G. (2008). The effect of an information leaflet upon knowledge and awareness of COPD in potential sufferers. *Respiration*, *76*(1), 53–60. doi:10.1159/000115947.

Cialdini, R.B. (2009). *Influence*. HarperCollins.

Cornish, D.B., & Clarke, R.V. (2003). Opportunities, precipitators and criminal decisions: A reply to Wortley's critique of situational crime prevention. *Crime Prevention Studies*, *16*, 41–96.

Craik, F., & Blankstein, K. (1975). Psychophysiology and human memory. In R. (Ed.), *In psychophysiology* (pp. 388–417). Wiley: London.

Cross, J. (2011). *Social Engineering is Often Overlooked*. Retrieved 23-October-2013, from http://www.immense.net/social-engineering-planning/.

Doob, A.N., & Gross, A.E. (1968). Status of frustrator as an inhibitor of Horn-Honking responses. *The Journal of Social Psychology*, *76*(2), 213–218. doi:10.1080/00224545.1968.9933615.

Ershoff, D.H., Mullen, P.D., Quinn, V.P. (1989). A randomized trial of a serialized self-help smoking cessation program for pregnant women in an HMO. *American Journal of Public Health*, *79*(2), 182–187. doi:10.2105/AJPH.79.2.182.

Ferguson, A.J. (2005). Fostering e-mail security awareness: the west point Carronade. *EDUCASE Quart*, *1*, 54–57.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford University Press.

Flight, I., Wilson, C., McGillivray, J. (2012). *Turning intention into behaviour: the effect of providing cues to action on participation rates for colorectal cancer screening. Colorectal Cancer-From Prevention to Patient Care*. Shanghai: InTech.

Ghaderi, F., Adl, A., Ranjbar, Z. (2013). Effect of a leaflet given to parents on knowledge of tooth avulsion. *European Journal of Paediatric Dentistry : Official Journal of European Academy of Paediatric Dentistry*, *14*(1), 13–6.

Gisquet-Verrier, P., & Riccio, D.C. (2012). Memory reactivation effects independent of reconsolidation. *Learning & memory (Cold Spring Harbor, N.Y.)*, *19*(9), 401–9. doi:10.1101/lm.026054.112.

Glanz, K., Rimer, B.K., National Cancer Institute, U. (1997). *Theory at a glance: a guide for health promotion practice.* U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health, National Cancer Institute.

Greenspan, S. (2008). *Annals of gullibility: why we get duped and how to avoid It.* Praeger.

Grewal, D., & Kavanoor, S. (1997). Comparative versus noncomparative advertising: a meta-analysis. *Journal of Marketing*, *61*(4), 1. doi:10.2307/1252083.

Gulas, C.S., & Weinberger, M.G. (2006). *Humor in advertising: a comprehensive analysis*. M.E. Sharpe, Incorporated.

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

Harris, P., Middleton, W., Joiner, R. (2000). The typical student as an in-group member: eliminating optimistic bias by reducing social distance. *European Journal of Social Psychology*, *30*(2), 235–253. doi:10.1002/(SICI)1099-0992.

Hart, A.R., Barone, T.L., Gay, S.P., Inglis, A., Griffin, L., Tallon, C.A., Mayberry, J.F. (1997). The effect on compliance of a health education leaflet in colorectal cancer screening in general practice in central England. *Journal of Epidemiology & Community Health*, *51*(2), 187–191. doi:10.1136/jech.51.2.187.

Hawkey, G.M., & Hawkey, C.J. (1989). Effect of information leaflets on knowledge in patients with gastrointestinal diseases. *Gut*, *30*(11), 1641–1646. doi:10.1136/gut.30.11.1641.

Hight, S.D. (2005). The importance of a security, education, training and awareness program. Retrieved 23-Oktober-2013, from http://www.infosecwriters.com/text_resources/pdf/SETA_SHight.pdf.

Hofstede, G., Hofstede, G.J., Minkov, M. (2010). *Cultures and organizations: software of the mind,* 3rd Edn. McGraw-Hill.

Humphris, G.M., Duncalf, M., Holt, D., Field, E. (1999). The experimental evaluation of an oral cancer information leaflet. *Oral Oncology*, *35*(6), 575–582. doi:10.1016/S1368-8375(99)00040-8.

Humphris, G.M., Ireland, R.S., Field, E.A. (2001). Randomised trial of the psychological effect of information about oral cancer in primary care settings. *Oral Oncology*, *37*(7), 548–552. doi:10.1016/S1368-8375(01)00017-3.

Krawczyk, A., Lau, E., Perez, S., Delisle, V., Amsel, R., Rosberger, Z. (2012). How to inform: comparing written and video education interventions to increase human papillomavirus knowledge and vaccination intentions in young adults. *Journal of American College Health : J of ACH*, *60*(4), 316–22. doi:10.1080/07448481.2011.615355.

Kumaraguru, P., Sheng, S., Acquisti, A., Cranor, L.F., Hong, J. (2010). Teaching Johnny not to fall for phish. *ACM Transactions on Internet Technology*, *10*(2), 1–31. doi:10.1145/1754393.1754396.

Lancaster, T., & Stead, L.F. (2005). *Self-help interventions for smoking cessation. Cochrane Database of Systematic Reviews*, *3*(3), CD001118. doi:10.1002/14651858.CD001118.

Lefkowitz, M., Blake, R.R., Mouton, J.S. (1955). Status factors in pedestrian violation of traffic signals. *The Journal of Abnormal and Social Psychology*, *51*(3), 704–706. doi:10.1037/h0042000.

Lien, N.H. (2001). Elaboration likelihood model in consumer research: a review. *Proceedings of the National Science Council*, *11*(4), 301–310.

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

Milgram, S. (1963). Behavioral study of obedience. *The Journal of Abnormal and Social Psychology*, *67*(4), 371–378. doi:10.1037/h0040525.

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

Mitnick, K.D., & Simon, W.L. (2002). *The art of deception: controlling the human element of security* Wiley.

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

Packer, D.J. (2008). Identifying systematic disobedience in milgram's Obedience experiments: a meta-analytic review. *Perspectives on Psychological Science*, *3*(4), 301–304. doi:10.1111/j.1745-6924.2008.00080.x.

Pallant, J. (2010). *SPSS Survival Manual: a step by step guide to data analysis using SPSS.* McGraw-Hill Education.

Petty, R.E., & Cacioppo, J.T. (1981). *Attitudes and Persuasion–classic and contemporary approaches.* W.C. Brown Company Publishers.

Petty, R.E., & Cacioppo, J.T. (1984). Source factors and the elaboration likelihood model of persuasion. *Advances in Consumer Research*, *11*(1), 668–672.

Petty, R.E., & Cacioppo, J.T. (1986). The elaboration likelihood model of persuasion. In Communication and persuasion (pp. 1–24). Springer.

Robb, K.A., Miles, A., Campbell, J., Evans, P., Wardle, J. (2006). Can cancer risk information raise awareness without increasing anxiety? A randomized trial. *Preventive Medicine*, *43*(3), 187–190. doi:10.1016/j.ypmed.2006.04.015.

Rogers, R.W. (1975). A Protection motivation theory of fear appeals and attitude change1. *The Journal of Psychology*, *91*(1), 93–114. doi:10.1080/00223980.1975.9915803.

Rosenstock, I.M. (1974). Historical Origins of the Health Belief Model. *Health Education & Behavior*, *2*(4), 328–335. doi:10.1177/109019817400200403.

Rouse, M. (2006). *Definition social engineering. TechTarget*. Retrieved 23-Oktober-2013, from http://www.searchsecurity.techtarget.com/definition/social-engineering.

Schellevis, J. (2011).*Grote Amerikaanse bedrijven vatbaar voor social engineering*. Retrieved 03-January-2014, from http://tweakers.net/nieuws/77755/grote-amerikaanse-bedrijven-vatbaar-voor-social-engineering.html.

Schmidt, S.R. (1994). Effects of humor on sentence memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 953.

Schneier, B. (2005). *Flaw in Winkhaus blue chip lock*. Retrieved 12-November-2013, from https://www.schneier.com/blog/archives/2005/03/flaw_in_winkhau.html.

Shim, S.M., Seo, S.H., Lee, Y., Moon, G.I., Kim, M.S., Park, J.H. (2011). Consumers' knowledge and safety perceptions of food additives: evaluation on the effectiveness of transmitting information on preservatives. *Food Control*, *22*(7), 1054–1060. doi:10.1016/j.foodcont.2011.01.001.

Stubbings, S., Robb, K., Waller, J., Ramirez, A., Austoker, J., Macleod, U., Hiom, S., Wardle, J. (2000). Development of a measurement tool to assess public awareness of cancer. *British Journal of Cancer*, *101*(S2), S13–S17. doi:10.1038/sj.bjc.6605385.

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

Weinstein, N.D. (1980). Unrealistic optimism about future life events. *Journal of personality and social psychology*, *39*(5), 806. doi:10.1037/0022-3514.39.5.806.

Jan-Willem H. Bullée is PhD-candidate at the Department of Services, Cyber-Security and Safety at the University of Twente, The Netherlands. His research interests include the psychological aspects of information security and the effect of interventions in information security.

Lorena Montoya is senior researcher at the Department of Services, Cyber-Security and Safety at the University of Twente, The Netherlands. Her research interests include crime science in general and in particular information security and spatio-temporal analysis of crime.

Wolter Pieters is technical leader of the European TREsPASS project on behalf of the University of Twente, and assistant professor cyber risk at Delft University of Technology. He has published widely on cyber security risk management, security policies and metrics, electronic voting, and philosophy and ethics of cyber security.

Marianne Junger is professor in cyber-security and business at the Department of Industrial Engineering and Business Information Systems at the University of Twente, The Netherlands. Her research interests include information security, cyber-crime science, and crime science.

Pieter Hartel is professor of computer science at the Department of Services, Cyber-Security and Safety at the University of Twente, The Netherlands. His research interests include information security and cyber-crime science.