Development of Cognitive Functioning Psychological Measures for the SEADM

F. Mouton, M.M. Malan, H.S. Venter

CSIR (Defence, Peace, Safety & Security), Command, Control & Information Warfare, Pretoria, South Africa
University of Pretoria, Information and Computer Security Architecture Research Group, Pretoria, South Africa
e-mail: moutonf@gmail.com

Abstract

Social engineering is a real threat to industries in this day and age, even though its severity is extremely downplayed. The difficulty with social engineering attacks is mostly the ability to identify them. Social engineers often target call centre employees, as they are normally underpaid, under-skilled and have limited knowledge about the information technology infrastructure. These employees are, thus, seen as easy targets by the social engineer. This paper improves on a previously-proposed model, Social Engineering Attack Detection Model (SEADM), by proposing and incorporating a cognitive functioning psychological measure in order to determine the emotional state and decision-making ability of the call centre employee. The cognitive analysis combined with the social engineering attack detection model provides one with a quick and effective way to determine whether the requester is trying to manipulate an individual into disclosing information for which the requester does not have authorization.

Keywords

Cognitive analysis, decision making, emotional state, SEADM, social engineering, Social Engineering Attack Detection Model, social psychology, information sensitivity.

1. Introduction

Social engineering, in the context of this paper, refers to various techniques that are utilised to obtain information through the exploitation of human vulnerability in order to bypass security systems (Mitnick & Simon, 2002). As clearly stated by various authors, the human element is the 'glitch' or vulnerable element within security systems (Scheeres, et al., 2008), (Mitnick & Simon, 2005), (Debrosse & Harley, 2009). It is the basic 'good' human-natured characteristics that make people vulnerable to the techniques used by social engineers, as it activates various psychological vulnerabilities that could be used to manipulate the individual into disclosing the requested information (Orgill, et al., 2004).

Individuals make themselves even more vulnerable to social engineering attacks by not expecting to ever be a victim of such an attack. Many may never even know that they were a victim of such an attack. The majority of the public, thus, may not fully

comprehend the extent to which these techniques to obtain such information, can be used. They also do not know the potential it holds for dire personal, economic and social consequences and losses for the individual as well as the institution. An individual may believe that the information they possess are of no particular value to another person, nor that it can be used for a malicious act. They may thus be more willing to disclose information freely. However, the social engineer is dedicated to researching various aspects and gathering information from various sources. Combined, the acquired information can have dire consequences.

On the other end of the spectrum, the individual may believe that they will not fall prey to such an attack, as they will be able to recognise such an attack. However, the social engineer is a skilled human manipulator, preying on human vulnerabilities using various psychological triggers that could foil human judgment.

The problem is to successfully detect social engineering attacks whilst working in a stressful environment, where decisions must be made instantaneously and under pressure. It is for this reason that the previously-proposed social engineering attack detection model (SEADM) by Bezuidenhout, Mouton and Venter (2010), has been improved upon by proposing a procedure in order to perform a cognitive functioning psychological measure. This cognitive functioning psychological measure is used to determine whether there is a change in the emotional state of the individual. It is also recommended that the model should be used in combination with training on various social engineering techniques, the psychological vulnerabilities it may elicit, and on institutional policies and procedures.

This research improves the SEADM by combining the two main perspectives of social engineering: the psychological perspective, and the computer science perspective. The psychological perspective focuses on the emotional state and cognitive abilities of the individual, whereas the computer science perspective focuses on information sensitivity. In essence, this research extends the existing SEADM with the addition of a cognitive functioning psychological measure.

The remainder of the paper is structured as follows. Section 2 provides background about social engineering. Section 3 introduces the previously-proposed SEADM in order to provide the reader with background knowledge of the original model. Section 4 proposes a new psychological measure to incorporate into the model by discussing the aim, the content and the results of such a measure. Section 5 provides an explanation of how the psychological measure should be incorporated into the model and what additional advantages it has to the previously-proposed model. Finally, Section 6 concludes with a summary on how the social engineering attack detection model has been improved on and provides suggested future work.

2. Social engineering

According to Mitnick & Simon (2002), social engineering is defined as the techniques used to exploit human vulnerability to bypass security systems in order to gather information. As indicated by this definition, social engineering attacks imply interaction with other individuals, indicating the psychological aspect of social engineering.

Various psychological vulnerabilities and triggers, used by social engineers, have been identified, which aim to influence the individual's emotional state and cognitive abilities in order to obtain information. To successfully defend against these psychological triggers, the individual will need to have a clear understanding of these triggers in order to recognise each during a social engineering attack. There are several psychological vulnerabilities, the most common ones are defined as: strong affect, overloading, reciprocation, diffusion of responsibility and moral duty, integrity and consistency, authority and finally deceptive relationships (Mitnick & Simon, 2005), (Gragg, 2002), (Workman, 2008), (Chantler & Broadhurst, 2006).

These triggers could be used to perform a social engineering attack on an unsuspecting victim, which could lead the victim to experience a sense of discomfort, whether just an uneasiness or even anxiety, as all these attacks prey on the victim's psychological vulnerabilities. One would expect that a victim would be able to use these clues of discomfort to detect that he is being targeted by a social engineering attack. However, this is the ideal and not reality, as the human reasoning and decision-making process is extremely complex, and prone to error.

The following section provides the practical application model, SEADM, which is used to determine whether a social engineering attack is being performed.

3. Social Engineering Attack Detection Model (SEADM)

In previous research, the authors have already proposed a social engineering attack detection model (SEADM). This model makes use of a decision tree by breaking the process down into more manageable components and guidelines to aid decision making. Figure 1 provides a shortened version of the SEADM, which consists of the two decision states which are focused on in this paper.

This model is used as a baseline throughout this paper and will improve on the parts where the individual is required to describe his or her own emotional state or provide their experienced level of discomfort. Throughout the remainder of the paper, the term *individual* is defined as the person dealing with the incoming call, as this model is proposed to be deployed within a call centre environment.

The first necessary step in this model would be for the individual to be conscious of, and able to evaluate, their emotional state on an ongoing basis. This implies a consciousness of emotion and how it can affect the individual's decisions.

In the same manner, the individual should evaluate the emotions that the person responsible for initiating the incoming call elicit within themselves, as the psychological vulnerabilities that might be triggered by a social engineering attack is directly aimed at creating certain emotional states, in order to obtain information.

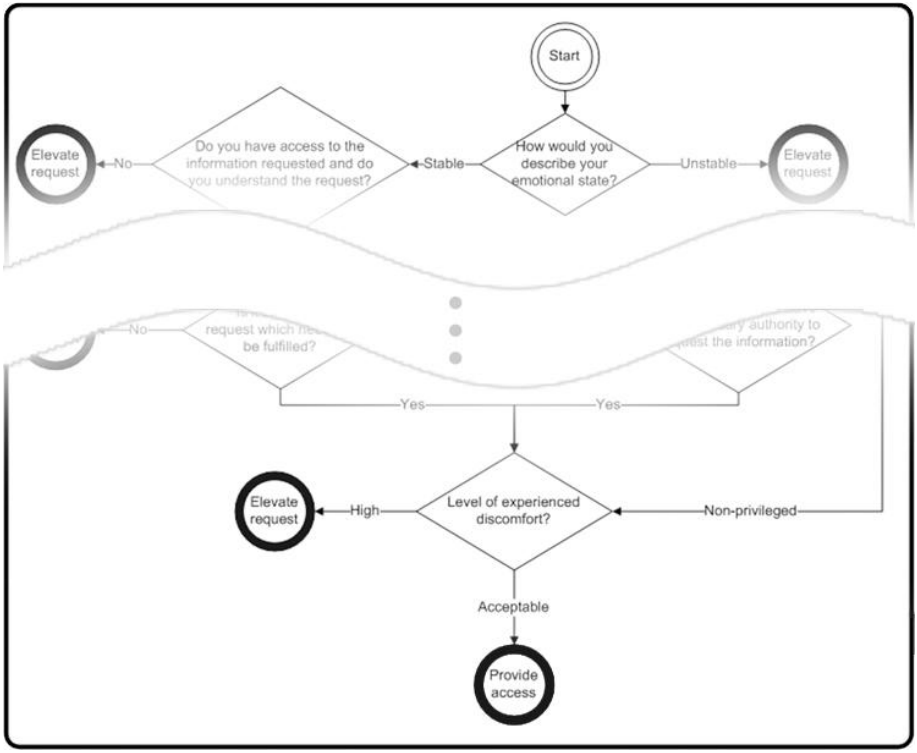

Figure 1: Social Engineering Attack Detection Model

One way or another we are all familiar with experiencing a bad day. Typically, such a bad day that starts off with some kind of bad experience and may seem to continue in such fashion throughout the day. For example, the car may break down on the way to work, followed by a negative emotional experience whether it results from family problems or an argument with a spouse or colleague. All factors and negative events influence our emotional state and hamper our ability to make rational, thought-through decisions (Siponen, 2008). Once a person finds themself in such a negative emotional state, that person is more likely to be a victim of social engineering: one's level of concentration may be low whilst irritability and frustration levels may typically be high, in which case an individual may be willing to provide a requester with certain information they rationally would have withheld, simply in a bid to get rid of the requester.

It is necessary to emphasise again the critical role an individual's emotional state can play in the safekeeping of privileged information. If an individual is in a negative

emotional state, the individual will not always be able to make a rational decision on the sensitivity level of the information of a request, or to whom it may be disclosed. This can result in costly losses to the institution and the individual.

The following section focuses on the development of a psychological measure which can be used to determine an individual's emotional state.

4. Developing a psychological measure for the SEADM

Awareness and consciousness of one's emotional state is not an easy task, or even always a possible task, for individuals. It is important to note that judging one's own emotional state could be a difficult task and some individuals are unable to perform this task in a rational way when their emotions are irrationally challenged. It is for this reason that an automated psychological measure is required.

The initial steps for the development of a psychological measure are that one clearly has to specify the aim of the measure, identify the content of the measure and analysing the assessment of the measure. These steps are discussed in the following respective subsections.

4.1. Specifying the aim of the psychological measure

It can be determined from the SEADM, as seen in figure 1, that one would need to determine both the emotional state and the level of experienced discomfort of the individual. This is a tedious task for individuals to determine their own emotional state, as each individual have their own interpretation and perception of emotional states as well as their level of experienced discomfort.

Initially, it is required to determine how one would go about determining both the emotional state and the level of experienced discomfort of an individual. It is important to note that the level of experienced discomfort is only a single attribute, which exists within an individual's emotional state. For this reason, if one is able to correctly determine the emotional state of an individual, this emotional state will encompass the level of experienced discomfort of an individual.

There are psychological measures which have already been developed to accurately determine one's emotional state and the level of experienced discomfort (Lopes, et al., 2003). These psychological measures which are able to determine the emotional state of an individual mostly comprise personality-based tests. The results of such personality-based tests can clearly indicate the emotional state of well-being of an individual (Lopes, et al., 2003). However, the major drawback of these personality tests is that they are lengthy and too cumbersome to be incorporated within this model.

Another issue to consider whilst assessing one's emotional state is that the emotional state of an individual is something which can stay constant for long periods of time. The emotional state of an individual only illicit a change when the individual is

exposed to issues such as experiencing severe stress, has a major life crisis or when the emotional well-being of the individual is affected by a health issue like illness.

It is for these reasons that it has been deemed impractical to assess the emotional state and level of experienced discomfort of the individual, by means of a psychological measure based on personality testing. In literature, it has been shown that there is a direct link between the performance of an individual on cognitive functioning based tests and the emotional state of an individual (Mathews, 1990), (MacLeod & Mathews, 1991).

According to Mathews (1990), "in any specific emotional state the cognitive system is organized in a manner appropriate for dealing with the new set of priorities arising from the particular event." As shown previously, if an individual is under attack by a social engineer, their stress levels would rise due to manipulation techniques exerted by the social engineer. Due to the increase in stress levels of the individual, the emotional state of the individual will change and, according to Mathews (1990), one can conclude that the change in emotional state will have an influence over the individual's cognitive functioning.

Social engineering attacks have also been shown to increase one's anxiety levels. This provides more evidence that any type of social engineering attack should have a direct influence over one's emotional state. This, in turn, will then have a direct influence over one's cognitive functioning.

Due to the intense effect that the emotional state has on the cognitive functioning of the individual, the aim of the psychological measure is to determine the level of cognitive functioning of an individual. The next section focuses on defining the content of the psychological measure.

4.2. Defining the content of the psychological measure

As it has been determined, by the authors, the aim of the psychological measure is to determine the level of cognitive functioning of an individual. This paper only examines pre-existing psychological measures which assess cognitive functioning. The field of cognitive functioning in psychology has been very well developed and several psychological measures have already been developed and are available for use (Eriksen, 1995), (Monchi, et al., 2001).

The advantage of using such psychological measures is that these psychological measures have already been approved and accepted by the American Psychological Association. The other advantage is that these psychological measures can also be analysed by a registered psychologist in order to determine the assessment results. It is very important to remember that, by law, one needs to be a registered psychologist in order to be allowed to draw conclusions and interpret the results of any psychological measure.

Whilst deciding which psychological measures to incorporate into the SEADM, it is important to remember that the entire model must be worked through during each

call taken by the individual. Only three psychological measures are considered for this paper due to space constraints, however, several other psychological measures can also be incorporated.

The three psychological measures that are considered can all be found in the Psychology Experiment Building Language (PEBL), which is an open source project that allows easy creation of computer based psychological measures (Mueller, 2012).

The three psychological measures, Wisconsin Card Sorting Test, Eriksen's Flanker Test and the Dot Judgement Test, which are discussed in the following subsections, are depicted in figure 2 (Eriksen, 1995), (Monchi, et al., 2001), (Cicchetti & Rourke, 2004).

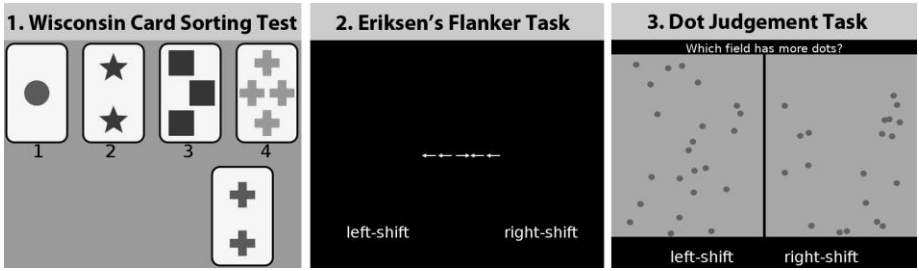

Figure 2: Psychological measures in PEBL

1. Wisconsin Card Sorting Test

The Wisconsin Card Sorting Test is normally used to test the ability to display flexibility in the face of changing schedules of reinforcement (Monchi, et al., 2001).

In the Wisconsin Card Sorting Test the individual is presented with a number of stimulus cards. The shapes on the cards are in different quantity, colour and design. The task of the individual is to match any additional cards by means of quantity, colour or design. The individual is then presented with a stack of cards in a specific order and is required to match each card to one of the stimulus cards. However, the individual is never told by which attribute the cards should be matched. The individual is only told if the match is correct or wrong. During the course of the test the rules of the game are changed. The time it then takes the individual to learn the new rules is measured as well as the amount of mistakes made in that time. The time taken and amount of mistakes which are made when the rules are altered, is used to determine the level of cognitive functioning of the individual.

2. Eriksen's Flanker Test

The Eriksen's Flanker Test assesses the ability to suppress responses that are incorrect in a specified context. Typically, in this test, a directional response is required from the individual when presented with stimuli which portrays conflicting or corresponding information about the directional response (Eriksen, 1995). In this

psychological measure, the individual is presented with five arrows of which the direction of the middle arrow is requested from the individual. If the middle arrow is pointing to the right-hand side, the individual is required to press the right-shift button. Similarly, when the middle arrow is pointing to the left-hand side, the individual is required to press the left-shift button. The arrows on either side of the middle arrow can provide either corresponding information, by pointing in the same direction as the middle arrow, or conflicting information, by pointing in an opposite direction. This psychological measure determines the time it takes for the individual to respond to the stimuli and the amount of errors which are made during the test.

3. Dot Judgement Task

The Dot Judgement Task assesses the space perception of an individual by requesting the individual to briefly count the amount of dots on a screen and indicate which of two sides has more dots (Cicchetti & Rourke, 2004).

In this psychological measure the individual is provided with two blocks containing red dots. This screen is only displayed to the individual momentarily. After the screen has been cleared, the individual is required to indicate which of the two sides of the screen has more dots by pressing either the left-shift button or the right-shift button. The psychological measure determines the time it takes for the individual to respond to each of the stimuli and the amount of errors made during the test.

All the psychological measures discussed previously are computer based tests and can be performed very briefly by an individual, and be assessed afterwards. The analysis of the assessment of the psychological measures is discussed in the following section.

4.3. Analysis of the assessment of the psychological measure

All of the psychological measures return numerical values, which need to be analysed before a conclusion can be drawn on the level of cognitive functioning of an individual. As mentioned earlier, the results of the assessments may only be interpreted by a registered psychologist. This obstacle can, however, be overcome as one knows that a specific state of cognitive functioning is indicated by the numerical values. Thus, an in-depth interpretation of the numerical values is not needed.

In order for the SEADM to function, it is required to determine whether there was a change in the emotional state of the individual. This can be determined by examining whether there was any change in the level of cognitive functioning of an individual, as the emotional state and cognitive performance influence each another. Determining if the emotional state of an individual has changed, requires only a comparison of the current level of cognitive functioning of the individual, versus the individual's normal level of cognitive functioning.

This comparison of cognitive functioning can be made using a feedforward neural network. Using a neural network allows one to determine if the emotional state of an individual has changed without performing a psychological analysis on the data. The

neural network can determine if the results of the psychological measure is significantly different from what is expected to be the constant level of cognitive functioning of the individual and, thus, the emotional state. This task of comparison is legally allowed to be done without the participation of a registered psychologist, as only the neural network interprets the data and detect changes in the level of cognitive functioning of the individual.

The feedforward neural network is an artificial neural network which uses a set of inputs with an associated output in order to correctly predict the outcome of any other similar input (Bebis & Georgiopoulos, 1994). The neural network is trained over time during the normal office duties of the call centre agent. Each new call provides one with new training data for the neural network for the specific individual. During each call, the individual is required to perform a psychological measure at both the start and the end of the call. On the initial step of evaluating one's emotional state, the psychological measure is used to determine if the level of cognitive functioning correlates to the individual's normal level of cognitive functioning and then returns a value, which is fed into the SEADM. The final step of the model, determining the level of discomfort, is performed by assessing whether the individual's level of cognitive functioning has been influenced during the call. In the case where the individual's level of cognitive functioning changes at any time between the start and the end of the SEADM, the individual would be deemed unfit to continue with the call, as the individual's emotional state would have been compromised. If the individual's emotional state has been compromised, the call is escalated to a more capable person whom can handle the call appropriately.

The training data for the neural network is collected continuously, whilst the individual is performing their normal office duties. This data is then used as a baseline when determining the individual's emotional state. The continuous updating of the training data allows the neural network to adapt to certain life conditions or life events which may occur to the individual.

Each separate test can have its own independent neural network associated to it, which is able to provide an output result. All the tests can also have a combined neural network which provides another independent output result. Having two independent sources, which can indicate if the emotional state has been compromised, allows one to have more accuracy and less false positives from the neural network. This is because both of the neural networks provide an indication whether the emotional state of the individual has been compromised or not.

The following section is devoted to incorporating the suggested psychological measures into the SEADM, as well as how to keep the assessments as short as possible whilst providing feasible results.

5. Incorporating psychological measures into the SEADM

It has been shown that incorporating cognitive functioning, psychological measures are an effective way to determine the level of an individual's emotional state. Both the initial state, where one has to evaluate one's own emotional state, and the state

where one has to evaluate one's level of experienced discomfort are replaced with incorporating the cognitive functioning psychological measure in the SEADM.

The individual is required to perform two, shortened versions from the group of the psychological measures, one at the initial state and one at the end state of the SEADM. The psychological measures provided to the individual only takes up to a maximum time span of thirty seconds, per measure, so that the efficiency in the call centre environment is not compromised. The shortening of the psychological measures has no adverse effects on the end result of the assessment, as the individual will complete several of these short versions during their normal office duties. This leads to an increase of one minute per call, thirty seconds at the start and thirty seconds at the end. It is a small price to pay to ensure that the individual in a call centre environment do not divulge any sensitive information, leading to a potential social engineering attack.

Randomising the order of the psychological measures reduce the repetitiveness of the tasks at hand. Using several different psychological measures has the advantage that the individuals will not get bored of having to complete the same psychological measure over and over again. This will also ensure that the user is presented with a new challenge each time. The order in which these psychological measures are completed is not important as each test has an independent neural network which is specifically trained on the individual associated with it.

It is, however, important to remember that the automated testing will only become effective when the individual has completed several attempts of each of the psychological measures. It is due to the design of the neural network that a large set of training data is required before it can be used to predict the output result, i.e. whether the individual's emotional state has been compromised. As the individual deals with more calls, the effectiveness of the automated testing improves as the neural network is provided with more training data. The initial training data can be acquired during an induction session directly after an individual is appointed. Collecting the training data during the induction session guarantees that there exists training data for each individual already. This ensures that when the individual starts his or her call centre agent duties, the neural network will already be trained to a sufficient level. During the initial training sessions, it is required that the individuals manually evaluate their own emotional states in order to provide training data to the neural network with the correct outputs. The accuracy of the neural network will be dependent on how accurately the neural network has been trained and on the amount of training data provided to the neural network.

The following section concludes the paper with a brief discussion of the improved SEADM and provides avenues for future research.

6. Conclusion

Social engineering is very difficult to detect, as the social engineer has various skills and effective techniques, preying on human vulnerabilities making these attacks often go without notice. What makes detection even more difficult is that many

people are unaware of this technique. Even though some people are aware of this technique, they might still not be aware of the potential threat and dire consequences it holds for the individual and for institutions.

It has been previously-proposed that a visible, practically applied, user-friendly aid, such as the SEADM, will aid in the daily awareness of the threat of social engineering attacks and, thus, protection against social engineering attacks. The purpose of this paper was to enhance the existing SEADM by changing the way individuals have to evaluate their own emotional state. This paper proposes to evaluate the individual's emotional state by making use of a cognitive functioning psychological measure using a feedforward neural network.

It has been shown that the emotional state of an individual is directly linked to the level of cognitive functioning of an individual. Due to this link between the emotional state and level of cognitive functioning, cognitive functioning psychological measures have been proposed. This psychological measure is able to determine if there has been any change in an individual's level of cognitive functioning. This change, in turn, would then indicate that the individual's emotional state has been compromised and that there may be a chance that the individual is being targeted by a social engineering attack. This improved model makes a valuable contribution to the field of social engineering, as it aids in the detection of social engineering attacks.

In further research one can also further explore research by Scheeres, Mills and Grimaila (2008) to illustrate the probable increase in awareness of an individual's own vulnerability to a social engineering attack through practical application of social engineering in a training environment. It would also be useful to perform some action research in a call centre in order to verify the usability of the improved SEADM.

7. Bibliography

Bebis, G. & Georgiopoulos, M., 1994. Feed-forward neural networks. *IEEE Potentials*, 13(4), pp. 27-31.

Bezuidenhout, M., Mouton, F. & Venter, H. S., 2010. *SEADM: Social Engineering Attack Detection Model.* Johannesburg, Information Security for South Africa (ISSA).

Chantler, A. N. & Broadhurst, R., 2006. *Social Engineering and Crime Prevention in Cyberspace,* Brisbane: Queensland University of Technology.

Cicchetti, D. V. & Rourke, B. P., 2004. *Methodological and Biostatistical Foundations of Clinical Neuropsychology and Medical and Health Disciplines.* 2nd ed. London: Taylor & Francis.

Debrosse, J. & Harley, D., 2009. *Malice through the looking glass: behaviour analysis for the next decade.* Geneva, Virus Bulletin Conference.

Eriksen, C. W., 1995. The flankers task and response competition: A useful tool for investigating a variety of cognitive problems. *Visual Cognition*, 2(2-3), pp. 101-118.

Gragg, D., 2002. A Multi-Layer Defense Against Social Engineering. *Sans Institute Reading Room*, December.

Lopes, P. N., Salovey, P. & Straus, R., 2003. Emotional intelligence, personality, and the perceived quality of social relationships. *Personality and Individual Differences*, 35(3), pp. 641-658.

MacLeod, C. & Mathews, A., 1991. Biased cognitive operations in anxiety: Accessibility of information or assignment of processing priorities?. *Behaviour Research and Therapy*, 29(6), pp. 599-610.

Mathews, A., 1990. Why worry? The cognitive function of anxiety. *Behaviour Research and Therapy*, 28(6), pp. 455-468.

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

Mitnick, K. D. & Simon, W. L., 2005. *The art of intrusion: the real stories behind the exploits of hackers, intruders and deceivers.* Indianapolis: Wiley Publishing.

Monchi, O. et al., 2001. Wisconsin Card Sorting Revisited: Distinct Neural Circuits Participating in Different Stages of the Task Identified by Event-Related Functional Magnetic Resonance Imaging. *The Journal of Neuroscience*, 21(19), pp. 7733-7741.

Mueller, S. T., 2012. *PEBL: The Psychology Experiment Building Language.* [Online] Available at: http://pebl.sourceforge.net/ [Accessed 13 January 2012].

Orgill, G. L., Romney, G. W., Bailey, M. G. & Orgill, P. M., 2004. *The urgency for effective user privacy-education to counter social engineering attacks on secure computer systems.* Salt Lake City, Information Technology Education.

Scheeres, J. W., Mills, R. F. & Grimaila, M. R., 2008. *Establishing the human firewall: reducing an individual's vulnerability to social engineering attacks.* s.l., 3rd International Conference on Information Warfare and Security.

Siponen, M. T., 2008. A conceptual foundation for organizational information security awareness. *Information Management & Computer Security*, 8(1), pp. 31-41.

Workman, M., 2008. A test of interventions for security threats from social engineering. *Information Management & Computer Security*, 16(5), pp. 463-483.