Considering the influence of human trust in practical social engineering exercises

WD Kearney
School of Computer, Statistical and Mathematical Sciences
North-West University
Potchefstroom, South Africa
Kearneys@iinet.net.au

HA Kruger
School of Computer, Statistical and Mathematical Sciences
North-West University
Potchefstroom, South Africa
Hennie.Kruger@nwu.ac.za

Abstract— **There are numerous technical advances in the field of information security. However, the application of information security technologies alone is often not sufficient to address security issues. Human factors play an increasing role in securing computer assets and are often detrimental to the security of an organisation. One of the salient aspects of security, which is linked to humans, is trust. It is safe to assume that trust will play an important role in any information security environment and may influence security behaviour significantly. In this paper the results of a practical phishing exercise and a trust survey are considered. The research project is part of a larger project and the phishing exercise is a follow-up to an earlier first practical phishing test. Results of the phishing test are compared with the first exercise. In addition, the newly obtained trust information from the survey is also incorporated into the report in order to try and explain security behaviour. The research was performed at a large organisation. Results indicate that although there is a general high level of trust in the organisation's ability to provide safe and secure information systems, a large number of staff was still victim to a simple phishing exercise. A possible explanation, which opens up further avenues for research, is offered.**

Keywords – Information security; Social engineering; Phishing;Trust

I. INTRODUCTION

Information security professionals know that users are often the weakest link in the information security chain. The famous hacker Kevin Mitnick had much success using social engineering – tricking people to give away sensitive data such as passwords [1]. There is a body of literature that shows technical controls work more effectively than the ability to manage the human aspects of information security. However, an important distinction that needs to be made is that technology is not the only answer in addressing information security risks, with attitudes and user perceptions playing an important part [2], [3].

More and more people are coming to realise that security failures are often due to issues other than the lack of suitable technical protection mechanisms. Some aspects are shown in the rapidly growing field of research in "Economics of Security" [4]. As part of this field, Moore and Anderson [5] describe active research with breaches of personal information and behavioural analysis.

The importance of addressing the human aspect in information security has grown over the past few years. One of the most frequent used techniques used to obtain private or confidential information from humans is phishing. Phishing is a kind of embezzlement that uses social engineering in order to obtain personal information from its victims, aiming to cause losses [6]. The Oxford English Dictionary [7] formally defines phishing as the fraudulent practice of sending e-mails purporting to be from reputable companies in order to induce individuals to reveal personal information, such as passwords and credit card numbers, online.

The Symantec Internet Security Threat Report [8] of April 2013 reported that e-mail phishing rates are down from one in 299 emails in 2011 to one in 414 in 2012. This does not, however, imply that the risk of being deceived has been reduced. The reason for this slight decrease is attributed to a shift in activity from email to social networks. Considering the billions of e-mail messages that are transmitted annually worldwide, it is clear that phishing attacks still form a considerable part of the day to day electronic communication activities and even with the slight decrease reported by Symantec, successful attacks may have a devastating effect on both enterprises and individuals. With this in mind it is safe to assume that technical as well as human controls become increasingly more important to mitigate or prevent phishing attacks.

It is also safe to assume that trust will play a significant role in any information security environment as good security will probably improve trust. Users' perceived security and perceived trust are closely related and it is therefore appropriate to consider human trust perceptions when dealing with social engineering and security awareness in general. There are many similar definitions of trust, the Macquarie Online Dictionary [9] describes trust as "on whom or that on which one relies" whilst another online dictionary definition states it as "a confidence that something is safe, reliable, or effective" [10]. The key words revolve around confidence and reliability. If one is confident that something is safe, reliable and effective, there would be a higher level of trust in that matter. Trust in this case refers to the human nature and not

978-1-4799-3383-9/14/$31.00 ©2014 IEEE

the computational notion of trust. It also refers, in this paper, to the sense of security or comfort a user has in the corporate environment, i.e. the level of confidence the user has in using the various systems.

This paper describes a practical social engineering experiment that was performed at a large organisation as a follow-up exercise to a previous practical exercise [11]. Apart from a mere comparison with previous results, a trust survey was also conducted to determine if trust has any influence in users' behaviour. The remainder of the paper is organised as follows. Section II presents, as background, a few examples of related research and also gives a very brief summary of the previous social engineering exercise. The methodology followed in this study is outlined in section III while section IV presents the results and a discussion of the current exercise. The paper is then concluded with general concluding remarks in section V.

II. Background and Related Work

A popular and effective way of addressing the human side of security is to focus on some form of an information security awareness program. Such a program can then concentrate on specific areas such as social engineering in general or phishing in particular. This is usually also an opportunity to emphasize the role of trust in an information security setup. There exist a large body of literature on these topics and the next few paragraphs will present some examples of such studies.

The acknowledgement that security breaches can be attributed to the behavior of computer users has led to a number of studies that were directed to users. Parsons *et al* [12] have developed a questionnaire to determine employee awareness by focusing on human aspects, while Crossler *et al* [13] highlighted directions for behavioral research in information security. Other examples of recent studies in this area can be found in [14] and [15].

Research studies on phishing, especially simulated attacks as reported on in this paper, were detailed in the first study of which this one is a follow-up and can be found in [11]. More recent examples can also be found in [6] and [16].

The possible role of trust forms an integral part of this paper and is consistent with other studies in this area. It is not unusual to find studies where trust is assessed in different systems or environments. Examples include trust in e-health systems [17], cloud computing [18], online purchasing [19] and e-payment systems [20].

As part of an ongoing study in understanding the management of information security risks, a first practical phishing exercise was conducted at a large geographically dispersed utility in 2012 and reported on in [11]. The organisation where the test was conducted is a large multi-billion dollar entity with over 3500 IT users and they supply essential services to over 2 million customers. During this first test, 280 users responded to a phishing message of whom 231 (83%) entered their usernames and passwords on a webpage. Of the 231 users, 23 (10%) entered their valid details more than once. A number of practical learning objectives were identified from the results of the first exercise. As part of this study, a follow-up practical test was undertaken together with a survey of users and management to assess their level of trust in the organisation's information systems.

III. Methodology

The methodology followed in this study comprised of two main steps. First, a questionnaire based survey was conducted to a broad spectrum of personnel to determine if any had been victims of a cybercrime, and also to establish whether those users had a level of trust in the corporation's ICT systems and infrastructure. This was then followed up by a practical e-mail based phishing exercise. The results of the phishing exercise were then evaluated and comparisons made to the original exercise [11] to determine if any change in behaviour had occurred or if any meaningful insights could be gained.

A. Trust survey

To gauge levels of trust and determine whether staff had been a victim of cybercrime before, a questionnaire was developed. The questionnaire consisted of 20 questions that were constructed based on management input and certain literature resources. The questions were specific to the organisation where the study was conducted and was tested with a small number of employees in a pilot run.

A sample of 40 users was used in the survey and included executive members, management and staff over a broad spectrum of the business. An appropriate sample size was difficult to determine as there were a myriad of factors that had to be taken into account, e.g. the sensitivity of the subject limited the sample size in this specific case. It was therefore decided to determine the sample size through a "saturation point" which is a standard stopping rule for research of this nature. Glaser and Strauss [21] used the term "theoretical saturation" which means that no additional data is found by the researcher for a specific category in a study. A disadvantage of this technique is of course that one would never know if new information can be obtained by questioning or interviewing an additional staff member. The same is however true for a statistically determined sample size. To ensure an appropriate response and to comply with the requirements of a saturation point stopping rule, the questionnaires were completed on an interview basis. An additional advantage of this approach was that the questions can be explained to respondents and in doing so ensure that all respondents understand the questions in the same manner. This hopefully increased the integrity of responses received.

Some of the questions had to be answered simply by indicating a yes or no. The objective of these questions was to establish a baseline e.g. whether users had been victims of cybercrime in the past 12 months. The majority of the questions had to be answered on a 5-point Likert scale and was aimed at assessing trust levels e.g. "*To what extend do you believe the Corporation provides a safe and trustworthy environment?*" There were also a few questions designed to deal primarily with the users' perception of whether they thought they had enough insight to both understand and manage their information risks. An example of such a question, which also had to be answered on a 5-point scale, is

the following. *"Do you have enough knowledge or information to manage your information risks?"* Interesting results were obtained from this first part of the study and will be presented in the section IV.

B. Phishing exercise

The practical phishing exercise implemented the same general and specific considerations used in the first exercise as discussed in detail in [11] except for a small change in the actual wording of the message. The structure and format of the e-mail was substantially similar but the message, whilst still relying on an explicit emotional exploit of scarcity, was modified to say: *"With our new password complexity rules, we require you to validate your username and password. If you act today, you will be in the draw to win a prize"*. One of the reasons for doing this was that the Symantec Internet Security Threat Report [8] stated that there is an increase of phishing scams that utilise fake websites and offer non-existent prizes.

This modification strengthened the legitimacy emotional exploit as the organisation where the exercise was conducted had recently modified their password complexity rules and length of password expiry as part of their ongoing information security risk management processes. This had been communicated throughout the organisation by poster, e-mail and articles in the in-house online magazine. The use of a prize was an added incentive.

To be able to perform a valid comparative analysis, the actual phishing exercise was conducted in the same manner as the first one and the same parameters were used. These parameters include sending out the message to all employees at 8:30 pm on a weekday night (the organisation is a 24-hour operation with activities taking place on a continuous basis). The reasons for this were the same as with the first exercise - to ensure that night workers are included in the test and to guarantee that day workers receive the message first thing in the morning. Following some concerns expressed with the first exercise, certain enhanced control measures were implemented, including ensuring the appropriate security personnel were notified. This follow-up actual test was allowed to run for an extended time. The extra time has provided further data for analysis which may provide further insight into the management of this important risk aspect. However, for the data analysis, only the dataset for the 12 hour test interval (the same as for the first exercise) were used.

Apart from the above specific issues, all general considerations to ensure the success of the project were also addressed, e.g. the obtaining of clearance and permission from the Chief Executive Officer to conduct the exercise, maintaining privacy of respondents etc.

IV. Results and Discussion

This section presents the results of the trust survey, the phishing exercise and comparative results with the initial first phishing experiment conducted in [11]. A possible explanation, especially with reference to the trust aspect, will also be presented.

A. Results of the trust survey

The overall result of the trust survey was clearly that there is a high level of trust amongst employees in the ability of the Corporation to provide a safe, secure and trustworthy environment. The following three questions are examples of evidence of this high level of trust that exist (all three were answered on a 5-point scale).

Q1: Do you think the Corporation protects and secures email communications and related data adequately? All respondents reacted by indicating that they believe that protection is adequate (either a 1 or a 2 on the 5-point scale – a 3 and higher indicates that they doubt the issue). This question is specifically significant to the email phishing exercise that was also conducted.

Q2: Do you feel confident enough in the corporate systems to do your online banking? The result was exactly the same as for the first question – all respondents feel confident to do online banking using corporate systems. This clearly implies a high level of trust in the corporate systems.

Q3: To what extent do you think the Corporate provides a secure or trustworthy IT environment? More than half, 57% rated it as a 1 (very secure) with 37% rated it as 2 (somewhat secure). Only 5% rated it as 3 (neutral) and no respondents rated it as 4 (not very secure) or 5 (very insecure). Figure 1 shows the results for this question graphically.

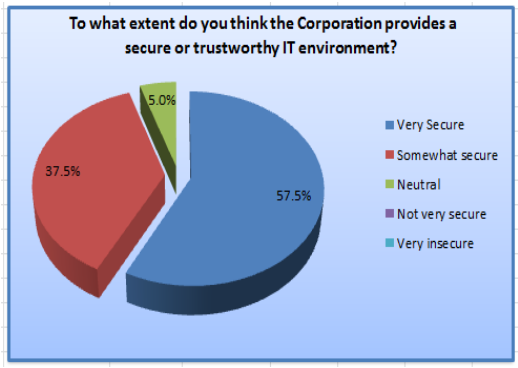

Figure 1. Secure and trustworthy environment

Another relevant question, mentioned in the methodology section, was *"Do you have enough knowledge or information to manage your information security risks?* The answers were somewhat illuminating in that only a small percentage of respondents believed they did not have enough information. Figure 2 graphically shows that over half were either somewhat (47.5%) or completely (12.5%) confident that they had enough knowledge to manage risks. The work of Schneier [22] shows that, on average, approximately 62% of employees have limited knowledge of information security risks whereas for this study 60% showed a positive slant – another indication of the high level of trust of employees in the Corporation and in their own security risk management capabilities.

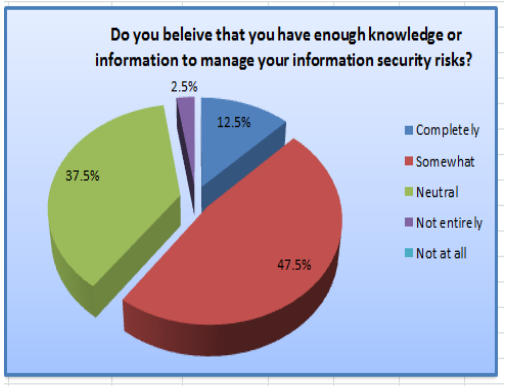

Figure 2. Knowledge to manage information risks

A final, interesting, remark on the trust survey is that more than half (55%) of the respondents had been a victim of any form of cybercrime in the past 12 months. This is notably higher than the reported 46% of computer user adults who had fallen victim of cybercrime in the last year [23].

B. Results of the phishing exercise

The same data as in the first test were recorded. This included user identification, section or department where the person works and time of access. Passwords were requested and were also validated through a technical process. To protect users' privacy, no passwords were recorded – only a simple yes or no was recorded depending on whether a valid password was entered or not. To ensure an acceptable level of data integrity, all duplicate records (users who entered their details more than once) and records with invalid usernames were removed from the final data set.

During the measured 12 hour time period of the test, 490 users responded to the phishing message of whom 312 (64%) entered their correct usernames and passwords on the webpage. A further 25 (5%) entered incorrect passwords and 154 (31%) users accessed the website but did not enter any credentials. A significant and somewhat disappointing statistic was the 48 users who accessed the website and who were repeat offenders in that they had entered their password correctly in the previous test. A total of 30 (63%) of these repeat offenders had entered their correct passwords again in the current exercise. Table 1 gives an overview of the statistics of users during the phishing exercise.

TABLE I. USER STATISTICS DURING THE PHISHING EXERCISE

Total employees	3500
Total number of users logged on for test	1400
Number of users who responded to the phishing message	490
Number of users who responded and who entered their passwords	312
Number of repeat offenders (first and current test)	48

It should be noted that although there were approximately 1400 active users logged on during the test, it would be incorrect to assume that all of those who did not respond recognised the phishing scam. There are certain reasons why many users did not respond to the phishing e-mail message. Some of the reasons include the fact that many people do not respond immediately to e-mail messages, others may have recognised the email for what it was and immediately deleted it, some users may have been engaged in other tasks and simply did not check their mail inboxes, etc.

One of the significant statistics computed during the first test was the number of users who entered their correct passwords and who has also completed a security training course. The objective of this security training course is to provide users with a basic level of security awareness so that they would be able to identify threats such as phishing scams. During the first test [11], 69% of the users who entered their correct passwords have also completed the security training. In this current test, the figure is very high at 92%.

The comparative results between the first test in [11] and this current follow-up test are summarised in table 2.

TABLE II. COMPARATIVE RESULTS OF THE TWO PHISHING EXERCISES

	First test [11]	Current follow-up test
Nr of responses	280	490
Nr of users who entered their user id's and passwords correctly	231 (83%)	312 (64%)
Nr of users who entered their passwords correctly and who previously completed security training	159 (69%)	288 (92%)

It is clear from table 2 that the results are quite unexpected and maybe somewhat disappointing. More people responded to the phishing e-mail message than in the initial exercise. Although there was a decrease in the percentage of users who entered their passwords, the physical number of users who did this increased by 81. The percentage number of users who completed the security training course and still gave away their passwords has also increased from 69% to 92%. This is an indication that the same concerns raised in the first exercise, still exist. These concerns are firstly, the high number of users who responded in a negative way despite their security training and secondly, the fact that there are still a number of users that never completed the compulsory information security course.

The results were also used to try and establish whether there is a link between experience (years of service) and being a victim of the phishing scam. In the first study it was found

that more than a third (35%) of those who entered their usernames and passwords has less than 5 years of experience and more than half (52%) had less than 10 years of experience. This was an indication that younger people (with less experience) are more prone to these types of security attacks. The results in this study confirmed this idea with just more than 50% of users who gave away their passwords having less than 5 years of experience, and a further more than 16% with less than 10 but more than 5 years of experience. These results are consistent with other research studies which focused on the same issues. Sheng *et al* [24], for example, used an online survey to try and determine who fell for phishing attacks. Their report shows that people aged 18-25 are more likely to be victims of a phishing scam when compared to the general population. Figure 3 shows an analysis of responses per experience category for this current study.

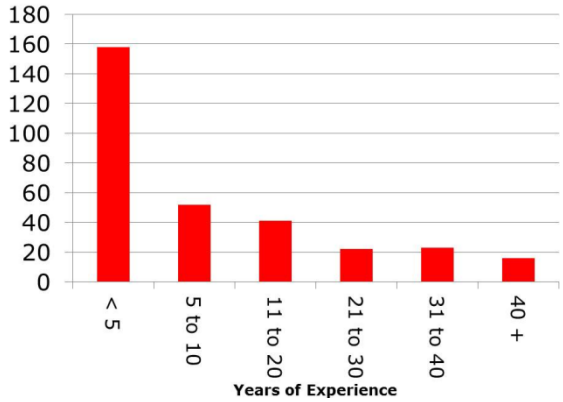

Figure 3. Responses per experience category

To summarize the results so far, it is clear that there is a high level of trust amongst users in their organisation's ability to provide a safe and secure information environment. Users also feel that they have on average enough skills and information to manage information security risks. However, a look at the results of the practical phishing test shows that a number of users became victims to the scam. It also appears as if they did not learn much from the previous test. The question now arises why this apparent contradiction in results? The next section will offer brief ideas or possible explanations for this situation.

C. Possible explanations

It seems like an anomaly when staff in an organisation fell, in large numbers, victim to a phishing scam while there is such a high level of trust in the organisation's information security environment. In addition, staff has indicated that they have sufficient knowledge to manage information security risks. The question arises why then do so many of them give away their passwords on web pages when asked for it?

The answer probably lies in the fact that information security is highly dependent on human factors. Aspects such as cognitive abilities, personal traits, perception of risk etc., plays a significant role and are most likely to impact security behavior. This case study has shown that the overwhelming majority of respondents have a positive perception of their own and their organisation's ability to protect them against security incidents such as phishing. High levels of trust seem to lead to carelessness where people are more easily tricked into security scams. It almost seems as if the level of trust impacts the level of risk behavior. This may be explained as follows.

Peltzman [25] put forward the concept of risk compensation in the safety arena to explain driver behaviour in adjusting individuals' levels of risk. Drivers would take more chances if they felt they were in a safer environment. Wilde [26] used this approach to develop the risk homeostasis theory. Risk homeostasis is based on the concept that people have a perceived or expected level of tolerable risk [27]. If there is a change in this level of risk they may compensate for it by changing their behaviour. For example, if the level of risk experienced by someone is low in comparison to the expected level of risk, he/she might engage in actions that will increase their exposure to risk. Conversely, if the level of experienced risk is higher than is acceptable, he/she may make an attempt to exercise greater caution.

This relates to information security in the sense that employees may become less vigilant or more careless when they know that good and adequate controls are implemented. E.g. users may become more easily victims of social engineering techniques such as phishing because they know (or perceive that) their organisation has the necessary controls (e.g. spam filters) in place? According to Pattinson and Anderson [27] there is not much doubt that risk homeostasis probably applies in many information security scenarios. They stated that risk homeostasis is after all a management theory and information security is all about managing risks. A similar link can be found in the medical field where some people believe that vaccinating young women against the human papillomavirus (HPV) will increase risky sexual behaviour [28] or in studies of sexual risk compensation such as in [29].

The trust survey conducted in this research has shown that users have a high level of trust in the Corporation's systems – at the same time a considerable number of employees fell victim to the phishing experiment. This seems to be in line with the above explanation of risk homeostasis. The level of risk experienced by users are low (results from the trust survey that indicates a high level of trust); users then compensate for this low risk by changing their behavior (taking more risks) and in so doing become phishing scam victims.

This paper forms part of a larger research project and in a next step of this larger project the role of risk homeostasis will be explored in more detail and reported on. It is hoped that the above theory on risk homeostasis will then be proved with more concrete examples and arguments.

V. CONCLUSION

With the acknowledgement that human factors play a significant role in the protection of information and information assets, the task of safeguarding these assets has

become more complex. To provide for risk perceptions, different attitudes and different levels of security knowledge is not an easy task. Criminals know that and focus their attacks on humans. A popular way of doing this is through social engineering attacks, more specifically phishing.

This paper forms part of a larger and ongoing project to investigate issues surrounding social engineering. In the first part of the project a practical phishing test was conducted at a large organisation. The results of this exercise were reported in [11]. In this current phase (this paper) a follow-up phishing test was performed at the same organisation. In addition, a trust survey was conducted to establish whether levels of trust may or may not play a role in being caught in a phishing scam. Interesting results were obtained. There was no real improvement in the number of people caught in the phishing scam; however, the trust survey revealed that respondents have a high level of trust in their own risk management abilities as well as in the ability of the organisation to provide them with a safe and secure information systems environment. No crystal clear explanation for this exist and the conclusion was that it is probably a case of risk homeostasis where users adjust their behavior (taking risks) to compensate for perceived low levels of existing risk (as indicated by the high level of trust).

There is already progress made with an ongoing research and investigation project that explores the risk homeostasis concept and applicability further. This is in an effort to gain more insight into the risk and security behaviour of people, especially in social engineering attacks.

REFERENCES

[1] K. Mitnick, The art of deception: Contolling the human element of security. Wiley, New York, 2002.

[2] S. Furnell and N. Clark, "Power to the people? The evolving recognition of human aspects of security," Computers and Security, 31, pp. 983-988, 2012.

[3] S. L. Pfleeger and D. D. Caputo, "Leverage behavioral science to mitigate cyber security risk," Computers and Security, 31, pp. 597-611, 2012.

[4] R. Anderson, "Economics and security resource," http://www.cl.cam.ac.uk/~rja14/econsec.html, Acessed: March 2014.

[5] T, Moore and R. Anderson, "Economics and internet security: A survey of recent analytical, empirical and behavioral research in internet security," in The Oxford handbook of the digital economy, M. Peitz and J. Waldfogel, Eds. Oxford University Press, 2011.

[6] C. K. Olivo, A. O. Santin and L. S. Oliveira, "Obtaining the threat model for e-mail phishing," Applied soft computing, 13, pp. 4841-4848, 2013.

[7] Oxford Dictionary, http://oxforddictionaries.com/definition/english/phishing, Accessed: February 2014.

[8] Internet Threat Security Report. Symantec Corporation, vol. 18, April 2013.

[9] Macquarie Dictionary, http://www.macquariedictionary.com.au, Accessed: February 2014.

[10] Macmillan online dictionary, http://www.macmillandictionary.com/dictionary/british/trust_23, Accessed: February 2014.

[11] W. D. Kearney and H. A. Kruger, "Phishing and organisational learning," in SEC2013, IFIP AICT 405, L. J. Janczewski, H. Wolf and S. Shenoi, Eds., pp. 379-390, 2013.

[12] K. Parsons, A. McCormac, M. Butavicus, M. Pattinson and C. Jerram, "Determining employee awareness using the human aspects of information security questionnaire (HAIS-Q)," Computers and Security, in press, 2014.

[13] R. E. Crossler, A. C. Johnston, P. B. Lowry, Q. Hu, M. Warkentin and R. Baskerville, "Future directions for behavioral information security research," Computers and Security, 32, pp. 90-101, 2013.

[14] T. Sommestad, J. Hallberg, K. Lundholm and J. Bengtsson, "Variables influencing information security policy compliance. A systematic review of quantitative studies," Information Management and Computer Security, vol. 22(1), pp. 42-75, 2014.

[15] K. Rantos, K. Fysarakis and C. Manifavas, "How effective is your security awareness program? An evaluation methodology," Information Security Journal: A global perspective, 21, pp. 328-345, 2012.

[16] S. Furnell, "Still on the hook: The persistent problem of phishing," Computer Fraud and Security, pp. 7-12, October 2013.

[17] S. Bahtiyar and M. U. Caglayan, "Trust assessment of security for e-health systems," Electronic Commerce Research and Applications, in press, 2013.

[18] R. Bose, X. Luo and Y. Liu, "The roles of security and trust: Comparing cloud computing and banking," Procedia – Social and Behavioral Sciences, 73, pp. 30-34, 2013.

[19] P. McCole, E. Ramsey and J. Williams, "Trust considerations on attitudes towards online purchasing: The moderating effect of privacy and security concerns," Journal of Business Research, 63, pp. 1018-1024, 2010.

[20] C. Kim, W. Tao, N. Shin and K. Kim, "An empirical study of customers' perceptions of security and trust in e-payment systems," Electronic Commerce Research and Applications, 9, pp. 84-95, 2010.

[21] B. G. Glaser, and A. L. Strauss, The discovery of grounded theory: strategies for qualitative research. New York, 1967.

[22] B. Schneier, "Insider threat statistics," http://www.schneier.com/blog/archives/2005/12/insider_threat.html, Accessed: March 2014.

[23] Symantec, "2012 Norton Cybercrime Report," http://now-static.norton.com/now/en/pu/images/Promotions/2012/cybercrimeReport/2012_Norton_Cybercrime_Report_Master_FINAL_050912.pdf Accessed: March 2014.

[24] S. Sheng, M. Holbrook, P. Kumaraguru, L. F. Cranor and J. Downs, "Who fallsfor phish? A demographic analysis of phishing susceptibility and effectiveness of interventions," Proceedings of the 28th International Conference on Human Factors in Computing Systems, pp. 373-382, 2010.

[25] S. Peltzman, "The effects of automobile safety regulation," Journal of Political Economy, vol. 83(4), pp. 677-726, August 1975.

[26] G. J. S. Wilde, Target risk. PDE Publications, Toronto, Canada, 1994

[27] M. R. Pattinson and G. Anderson, "Risk homeostasis as a factor of information security," http://www.igneous.scis.ecu.edu.au, Accessed January 2014.

[28] N. T. Brewer, L. C. Cuite, J. E. Herrington and N. D. Weinstein, "Risk compensation and vaccination: Can getting vaccinated cause people to engage in risky behaviours?," Annals of Behavioural Medicine, vol. 34(1), pp. 95-98, 2007.

[29] S. D. Pinkerton, "Sexual risk compensation and HIV/STD transmission: Empirical evidence and theoretical considerations," Risk Analysis, vol. 21(4), pp. 727-736, 2001.