



A vocabulary test to assess information security awareness

Hennie Kruger, Lynette Drevin and Tjaart Steyn
*School of Computer, Statistical and Mathematical Sciences,
North-West University, Potchefstroom, South Africa*

Abstract

Purpose – The dependence on human involvement and human behavior to protect information assets necessitates an information security awareness program to make people aware of their roles and responsibilities towards information security. The purpose of this paper is to examine the feasibility of an information security vocabulary test as an aid to assess awareness levels and to assist with the identification of suitable areas or topics to be included in an information security awareness program.

Design/methodology/approach – A questionnaire has been designed to test and illustrate the feasibility of a vocabulary test. The questionnaire consists of two sections – a first section to perform a vocabulary test and a second one to evaluate respondents' behavior. Two different class groups of students at a university were used as a sample.

Findings – The research findings confirmed that the use of a vocabulary test to assess security awareness levels will be beneficial. A significant relationship between knowledge of concepts (vocabulary) and behavior was observed.

Originality/value – The paper introduces a new approach to evaluate people's information security awareness levels by employing an information security vocabulary test. This new approach can assist management to plan and evaluate interventions and to facilitate best practice in information security. Aspects of cognitive psychology and language were taken into account in this research project, indicating the interaction and influence between apparently different disciplines.

Keywords Data security, Semantics, Cognition, Language, Tests and testing

Paper type Research paper

1. Introduction

The aim of information security is to ensure business continuity and to minimize business damage by preventing and minimizing the impact of security incidents (von Solms, 1998). In general, information security refers to the following three important aspects (Pfleeger and Pfleeger, 2007):

- (1) *Confidentiality*. Computer-related assets are accessed by authorised parties only.
- (2) *Integrity*. Correctness of computer assets such as data; data cannot be modified by unauthorised parties.
- (3) *Availability*. Computer-related assets are accessible to authorised parties at appropriate times.

The protection of information assets usually relies on the success of information security plans and the implementation of various security controls as part of such a plan. Apart from the usual technical controls, there is also a huge dependence on human involvement, and this human factor in information security is directly related to human

This paper is based upon work that was presented at the South African Information Security Multi-conference in Port Elizabeth, South Africa, 17-18 May 2010 (Kruger *et al.*, 2010).



behavior and human knowledge. This means that humans involved in a security process need to possess the required knowledge about their security-related roles and thus need some form of education (van Niekerk, 2005).

To address this need for educating people and making them aware of information security threats, organizations often make use of information security awareness programs. According to Dhillon (1999), the user education, or awareness program, is singled out because increasing awareness of security issues is the most cost-effective control that an organization can implement. This implies that a certain financial investment is required to design and implement an information security awareness program. Such an investment can become significant and a well-lanned strategy is necessary to support the goals of an awareness campaign and to target those areas where specific needs exist. By understanding the various information security issues that might exist, it becomes possible to identify appropriate approaches that could be adopted to overcome information security awareness obstacles.

In this paper, a survey will be described which determines the feasibility of a vocabulary test to identify areas to focus on in an information security awareness program. The study is based on another study performed from an educational viewpoint where the mathematics vocabulary of school learners was evaluated in order to identify specific mathematics learning areas that may need special attention (van der Walt *et al.*, 2008). This research is an example of interdisciplinary research where theories of human cognitive development and language are taken into account when a seemingly unrelated discipline, e.g. information and communication technology security, is studied.

The approach to make use of techniques borrowed from other disciplines is not new and many researchers have already done this. For example, Maseti and Pottas (2006) investigated the applicability of a role-based information security awareness model in South African hospitals, while van Niekerk and von Solms (2004) examined the educational principles of outcomes-based education in an information security awareness program. With respect to information security culture, Schlienger and Teufel (2003) made use of social-cultural measures to define a model for analyzing information security culture in organizations.

The remainder of this paper is organized as follows. In Section 2, a brief theoretical background is given. The methodology followed is presented in Section 3 while Section 4 discusses the results of an empirical test. Section 5 concludes the paper with some final comments.

2. Theoretical background

The assessment of information security awareness levels in this study is completely based on an educational study performed to test school learners' mathematics vocabulary in order to plan and evaluate interventions and to facilitate best practice in mathematics classrooms (van der Walt *et al.*, 2008). In this educational study, a mathematics vocabulary questionnaire was developed to measure learners' language proficiency in mathematics. Based on the results, the researchers were not just able to predict future mathematics performance, but more importantly, they were also able to provide guidelines that can be used to identify specific focus areas to concentrate on in order to improve mathematics performance. This same approach and principles were applied to the information security awareness project described in this paper because the set of familiar words related to information security constitutes one's information

security vocabulary. Such a vocabulary will evolve over time and enables a person to communicate and acquire new knowledge.

To understand the complexities of language, vocabularies and associated processes of transforming inputs into beliefs and actions, it is necessary to briefly refer to a few aspects from the field of cognitive psychology.

Cognitive psychology is concerned with internal mental processes and behavior (Robinson-Riegler and Robinson-Riegler, 2008). The metaphor of a computer is often used to describe the cognitive psychology idea of how people acquire, process and store information (Robinson-Riegler and Robinson-Riegler, 2008; Slater and Bremner, 2005). The human mind is viewed as an information processing system in order to solve problems (Russel and Norvig, 1995) and the mental state then comprises of at least three components:

- (1) information is received (from the environment). This could be encoded in some form;
- (2) internal processes take place (memory storage, application of problem-solving methods, etc.); and
- (3) Cognitive structures of the individual are changed in reaction to the information.

Over time, changes may occur. First, the capacity of processing units changes and second, the effectiveness and variety of strategies improve.

One of the leading experts in the area of language studies, Noam Chomsky, explained the learning of language from a cognitive psychology angle and states that one is born with language abilities (Robinson-Riegler and Robinson-Riegler, 2008). This theory that language depends on rules that are within us and that it is not learnable is challenged by more recent theories that suggest that structural aspects of language are indeed learnable. This originates from a neural network approach to cognition which argues that learning takes place through associations in the neural network system throughout the human brain (Robinson-Riegler and Robinson-Riegler, 2008).

Slater and Bremner (2005) argued that there is an interaction between language and cognitive development and that children may have formed concepts of the world even before they acquire words. Children's first words will probably be formed from early concepts because their communication is based on what they know, and as semantic systems develop, so will children's concepts change. Hearing new words will also result in searching for the meaning of the word and thereby results in the learning of new words. When gaps are faced in the semantic or cognitive system, two strategies can be used (Slater and Bremner, 2005):

- (1) acquire a new word – search known concepts or construct a new one; and
- (2) acquire a new concept – attach a known word to it or look for one.

Language development (which is based on vocabularies) is complex and many theories tried to describe this phenomenon. A more detailed description of language development and vocabularies is beyond the scope of this paper and readers interested in this topic are referred to sources on developmental psychology and cognitive psychology such as those used in this paper.

According to van der Walt (2008), there are a number of factors that will have an impact on a person's study and learning orientation. These factors include cognitive (acquiring and implementation of knowledge) and meta-cognitive (monitoring and evaluation of cognitive strategies) factors, affective factors (how learning experiences influence people's perceptions), conative factors (translation of knowledge and emotion into behavior) and cross-cultural factors (the effect of social and environmental influences on behavior).

It is clear that all these factors will also play a role in how basic information security principles are learned and applied. In this study, however, the focus will only be on the cognitive aspect. This decision is based on the fact that the project is an exploratory study to assess the idea of applying a vocabulary test to information security awareness training. Once the use of cognitive principles has been established as being useful in the information security awareness arena, the investigation of the other factors would form part of a follow-up project to create a more comprehensive information security awareness assessment and training model.

A basic definition of the term cognitive is given by van der Walt (2008) as the process to learn, know, understand, code, process and recall information. An international report issued by Trends in International Mathematics and Science Study (Mullis *et al.*, 2005) states that there are three key cognitive skills necessary for a successful learning experience. They are:

- (1) knowledge of facts, processes and concepts;
- (2) the ability to apply the knowledge, concepts and processes; and
- (3) the ability to reason.

Table I contains an explanation of the three required cognitive skills as it pertains to information security awareness.

With this very brief theoretical background in mind and in line with the mathematical proficiency study, it was decided to investigate the feasibility of using an information security vocabulary test in order to be able to draw conclusions about people's information security awareness levels. The motivation lies in the fact that if one does not have a basic comprehension of information security concepts or terms, then one is probably more susceptible to become a victim of security attacks. In addition to this, it may render certain security awareness material useless, e.g. a poster warning you against phishing will not have the desired effect if you do not know what the term phishing means. Having said this, it should be noted that a vocabulary test on its own may produce misleading results, e.g. in some cases, one does not really have to know what a specific term means to be cautious to certain requests, e.g. you may decide that giving out personal details is wrong without knowing what the term phishing means. This problem was addressed with scenario-type questions to test respondents' behavior. The questionnaire containing both vocabulary tests and the behavior tests will be discussed in the next section.

3. Methodology

A questionnaire was developed to test and illustrate the feasibility of a vocabulary test. The questionnaire consists of two sections – a first section to perform a vocabulary test and a second one to evaluate respondents' behavior.

Cognitive category	Cognitive action	Explanation
Knowledge of facts, processes, procedures and concepts (what someone needs to know)	Recall, recognize, calculate, derive information from graphs or tables, measure, classify and sort	When people do not have reasonable access to a knowledge or facts base in information security, focused information security reasoning becomes difficult. Knowledge of security processes (steps, methods or procedures) forms the link between basic knowledge and the implementation thereof. Knowledge of information security concepts enable people to see the relationship among the different elements of information security and help to ensure that facts are not seen or treated in isolation
Understanding and application of knowledge	Choose, suggest, develop a model, solve problems and implement solutions	Representation of information security ideas forms the basis of perceptions and communication in information security and is a basic prerequisite for a successful information security environment. When it is expected from someone to apply knowledge in the information security area, the type of problem should be known in order to execute the required procedures and to choose the best strategy for solving the problem
Reasoning (focus on solving problems in unknown situations)	Analyze, generalize, integrate defend solutions	Reasoning in information security requires logical and systematically, including intuitive and inductive, thinking processes. People should be able to implement expertise in different contexts

Table I.
Cognitive skills

Source: Adapted from van der Walt (2008)

The idea with the first section, the vocabulary test, was to include basic and “generally known” concepts and terms to establish whether respondents know the meaning of these concepts. Because this study is an exploratory study to test the feasibility of a vocabulary test, it was decided to start with only eight of the more general security aspects. For the same reason, it was also decided not to include “less known” concepts such as botnets, steganography, etc. To ensure that the final eight concepts used in the questionnaire were relevant, the selection of the concepts was based on information in the 2008 Information

Security Breaches Survey conducted by PriceWaterhouseCoopers (2008) and a SANS (2005) Institute report on e-mail security threats. The questions were, consistent with the mathematics study, constructed as multiple choice questions with five options to choose from. As an example, Example 1 shows the vocabulary test for the term phishing – it is expected that someone with a good understanding of the term phishing would select option (d) as the correct or most appropriate answer.

Example 1. Vocabulary test

Phishing is:

- (a) The use of an e-mail message, that appears to be legitimate, to solicit personal details.
- (b) Part of social engineering which means that someone is persuaded to give away confidential information.
- (c) Also referred to as identity theft.
- (d) All of the above.
- (e) I do not know what the term phishing means.

Another example of a question from the vocabulary test is as follows.

Example 2. Vocabulary test

Spam is:

- (a) Another word for e-mail or electronic messages.
- (b) A marketing technique.
- (c) Any unsolicited electronic mail.
- (d) All of the above.
- (e) I do not know what spam is.

Section 2 of the questionnaire contained scenario-type questions to evaluate respondents' behavior independently of their vocabulary knowledge. The majority of the questions were (with small adjustments) taken from a security awareness index report prepared during 2002 by Pentasafe Security Technologies (2002). The work of Furnell *et al.* (2007) was also used to construct some of the questions. A total of nine questions were asked to test the behavior linked to some of the concepts in section 1 of the questionnaire. For example, for the word phishing in section 1 of the questionnaire, the following scenario type of question (Example) was used to test whether a respondent will withhold personal details irrespective of whether the respondent knows what the term phishing means.

Example 3. Scenario-type question

When receiving an e-mail that appears to be coming from your bank and asking you to go to a specific web link to confirm your personal details, what would you do? (select ALL that apply):

- (a) If the bank's logo, address and all other information on the e-mail and webpage are correct, I will provide the required information.
- (b) I will simply ignore the request.

- (c) If my colleagues received the same request and if they have provided their details, I will do the same.
- (d) I will phone the bank to find out about the request.
- (e) I will report it to our company's IT department.

The questionnaire was sent to two different class groups of students at a university and 44 responses were received. The reason for using students as a test base and for accepting the relatively low number of responses is because the objective of the study was purely to test the concept of using a vocabulary approach for information security awareness purposes.

This study will be followed up with a more comprehensive one where a bigger population of employees in the corporate environment will be targeted. For the same reason, no extensive statistical validations were performed to standardize the questionnaire – this will also form part of the follow-up study.

4. Results

4.1 Demographic details

From the responses received, 84 per cent were male and 16 per cent female. The major fields of study of the respondents are shown in Figure 1 where it can be seen that the majority studied in the natural sciences.

Figure 2 shows the year levels of the respondents. Most of the respondents (38.6 per cent) were fourth year or honors students.

4.2 Interpretation of data

Based on the results of the vocabulary test, it appears that the majority of respondents have a reasonable knowledge of threats linked to e-mail security such as computer virus, worm, spyware, spam, etc. With regard to phishing, it was quite surprising that 11 per cent indicated that they do not know what the term means while more than

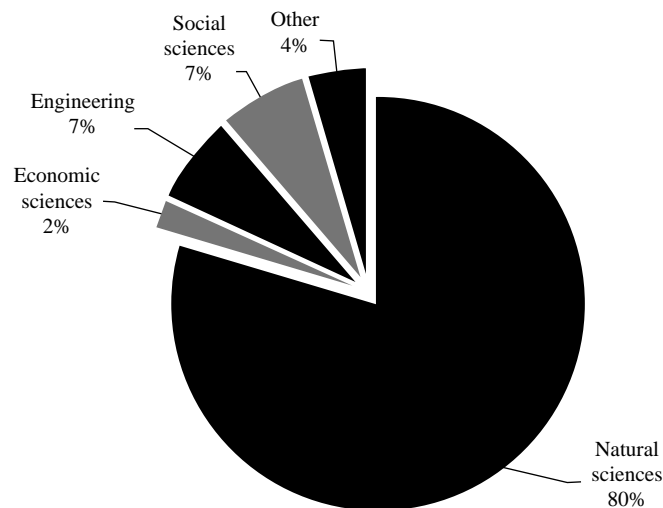


Figure 1.
Fields of study

25 per cent do not understand the term security incident. It was also clear from the scenario type of questions that respondents are not really sure where to report security incidents. Figure 3 shows the results to the question to whom would you first report a security incident. The figures are presented as percentages and it should be noted that they do not add up to 100 as respondents may have selected more than one alternative.

The most surprising result, however, was the fact that respondents do not have a good comprehension of one of the most basic security aspects. Almost half of them (48 per cent) did not know what the term strong password means. This inability to understand the strong password concept was strongly supported by the results of the behavior tests in the scenario-type questions related to passwords. Figure 4 shows that although the majority has indicated that they memorize their passwords, a significant number of respondents are still making use of unacceptable techniques to remember their passwords. The question was what techniques do you use to remember your password? (Respondents may have selected more than one alternative).

In addition to unacceptable techniques to remember passwords, there were also a number of respondents who were willing to give their passwords away under certain circumstances. Figure 5 shows the results (in percentages) to the question, to which of the following people would you tell your password if that person requested it (respondents may have selected more than one alternative).

Owing to paper length constraints, not all results can be shown or discussed here, and a last graph that shows how respondents feel about information security is shown in Figure 6. The notion is that if one's vocabulary on a specific subject is not good,

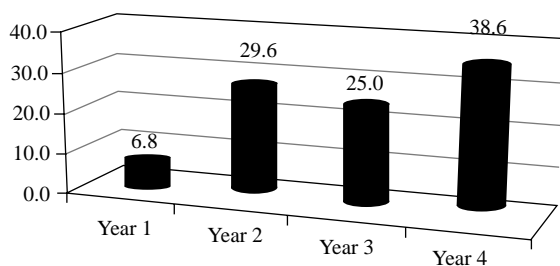


Figure 2. Year levels of respondents

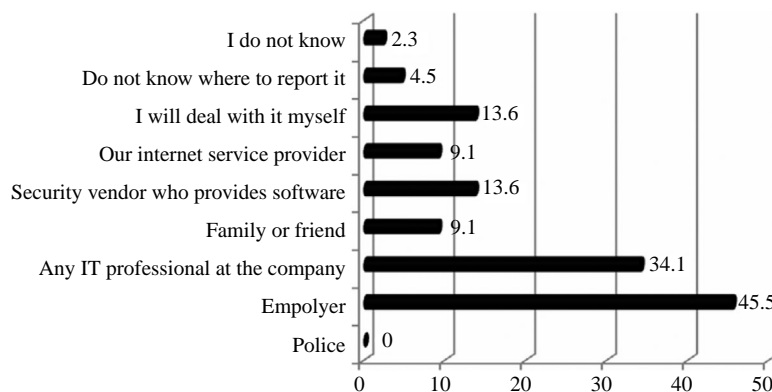


Figure 3. Reporting of security incidents

then one's attitude (how you feel) towards this subject is probably not too positive. Respondents simply had to select all statements on how they feel about information security. The figures are given as percentages.

The results show that a vocabulary test can make a definite contribution to identify specific areas for security education. In the case of this study, it was clear that any security awareness program, for the students under discussion, should concentrate on two aspects, namely security incidents, e.g. what constitutes a security incident (knowledge) and where and how should it be reported (knowledge and behavior). Second, the awareness program should focus on strong passwords, e.g. what does the concept mean (knowledge) and what should be done (behavior) to ensure that a password is a strong password and that it remains a strong password. Only a smaller portion of the awareness program should then be dedicated to aspects such as viruses and spam as the vocabulary test, and the scenario tests have shown that the students do understand these terms and that their behavior in this regard is satisfactory.

The overall objective of the study was to test the feasibility of using a vocabulary test to assess information security awareness. Based on the reported figures and graphs above, it is clear that the use of a vocabulary test will be beneficial. This conclusion is further supported by a simple and basic analysis of the relationship between knowledge of concepts (vocabulary) and corresponding behavior. For example, 89 per cent of respondents had a fairly good idea of what the term phishing means. This was supported by a positive corresponding behavior result where only 9 per cent indicated that they will provide their personal details if requested by e-mail. Another example

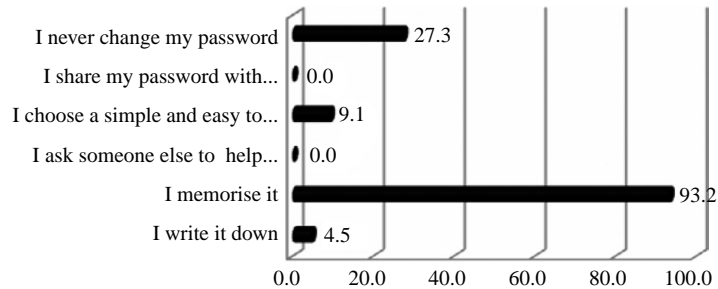


Figure 4.
Techniques to remember passwords

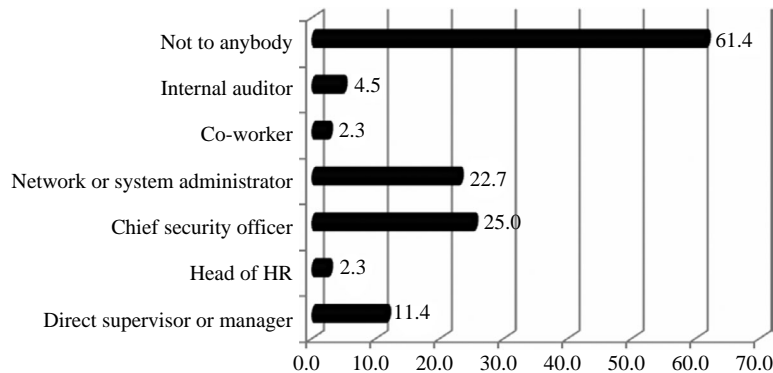


Figure 5.
To whom will you give your password

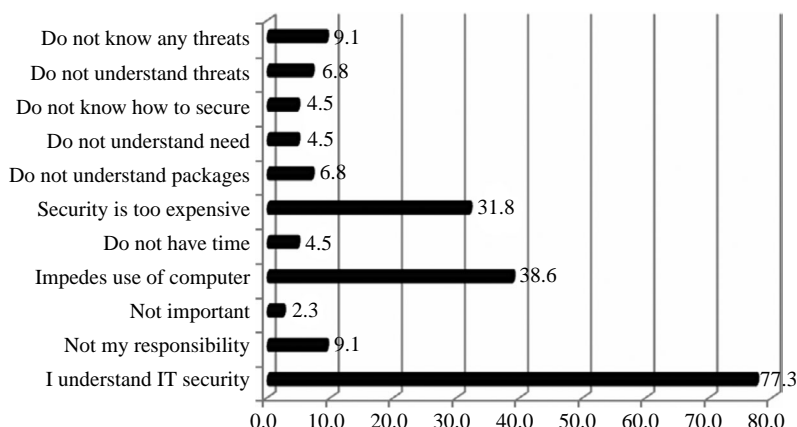


Figure 6.
How respondents feel about information security

of the correlation between vocabulary and behavior was the concept of a strong password mentioned earlier on. Almost half of the respondents did not know what it means. This lack of knowledge was confirmed with the behavior type of questions where a significant number of respondents have indicated that they make use of unacceptable techniques to remember their passwords and that some of them are willing to reveal their passwords on request. Other cases supporting the relationship between a vocabulary test and security awareness were also found but are not reported here.

The results of this study and the reported figures and data presentations in this paper are by no means a comprehensive report on the use of vocabulary tests in information security. Only a limited number of concepts were included in the vocabulary questionnaire and the sample size for the survey can also be expanded. This exploratory study and results obtained do, however, confirm that a vocabulary test can be useful to plan and evaluate interventions and to facilitate best practice in information security.

5. Conclusions

If modern organizations want to survive and prosper, it is vital that their employees have both the necessary knowledge and the right attitude to fulfil their required roles and responsibilities in the overall information security efforts of a company (van Niekerk, 2005). One way of addressing this need is to design and implement suitable information security awareness programs.

This paper described an exploratory study to test the feasibility of employing an information security vocabulary test to assist management in identifying specific areas or topics to be included in an information security awareness program. To test this concept, a questionnaire consisting of a vocabulary and a behavior test was developed and distributed to a small population of students. The results of the survey confirmed that an information security vocabulary test is useful and should be considered to assess information security awareness levels and to identify focus areas for intervention.

The intention is to expand this research project to develop a statistically validated vocabulary test and then to repeat the exercise in the corporate and other relevant environments. In addition, further investigation of cognitive psychology and language aspects will be conducted in order to align the development of security awareness vocabulary tests with human cognitive developmental issues.

References

- Dhillon, G. (1999), "Managing and controlling computer misuse", *Information Management & Computer Security*, Vol. 7 No. 4, pp. 171-5.
- Furnell, S.M., Bryant, P. and Phippen, A.D. (2007), "Assessing the security perceptions of personal internet users", *Computers and Security*, Vol. 26 No. 5, pp. 410-17.
- Kruger, H.A., Drevin, L. and Steyn, T. (2010), "The use of an information security vocabulary test to assess information security awareness – an exploratory study", in Clarke, N., Furnell, S. and von Solms, R. (Eds), *Proceedings of the South African Information Security Multi-Conference, Port Elizabeth, South Africa 17-18 May*, Centre for Security, Communications & Network Research, University of Plymouth, Plymouth.
- Maseti, O. and Pottas, D. (2006), "A role-based security awareness model for South African hospitals", *Proceedings of the 6th Annual Information Security South Africa Conference, Sandton, South Africa, 5-7 July*.
- Mullis, I.V.S., Martin, M.O. and Foy, P. (2005), *IEA's TIMMS 2003, International Report on Achievement in the Mathematics Cognitive Domains*, International Association for the Evaluation of Educational Achievement (IEA), Boston college, Boston, MA.
- Pentasec Security Technologies (2002), *Security Awareness Index Report: The State of Security Awareness among Organizations Worldwide*, Pentasec Security Technologies, Farnham.
- Pfleeger, C.P. and Pfleeger, S.L. (2007), *Security in Computing*, 4th ed., Prentice-Hall, Upper Saddle River, NJ.
- PriceWaterhouseCoopers (2008), "Information security breaches survey", Technical report, available at: www.security-survey.gov.uk (accessed 25 November 2009).
- Robinson-Riegler, G. and Robinson-Riegler, B. (2008), *Cognitive Psychology Applying the Science of Mind*, Pearson, Boston, MA.
- Russel, S. and Norvig, P. (1995), *Artificial Intelligence: A Modern Approach*, Prentice-Hall, Englewood Cliffs, NJ.
- SANS (2005), *E-mail Security Threats*, SANS Institute InfoSec Reading Room, Bethesda, MA.
- Schlienger, T. and Teufel, S. (2003), "Information security culture – from analysis to change", *South African Computer Journal*, Vol. 31, pp. 46-52.
- Slater, A. and Bremner, G. (2005), *An Introduction to Developmental Psychology*, Blackwell, Padstow.
- van der Walt, M.S. (2008), "Aanpassing van die studie oriëntasievraelys in Wiskunde vir gebruik in die intermediêre fase", PhD dissertation, North-West University, Potchefstroom.
- van der Walt, M.S., Maree, K. and Ellis, S. (2008), "A mathematics vocabulary questionnaire for use in the intermediate phase", *South African Journal of Education*, No. 28, pp. 489-504.
- van Niekerk, J.F. (2005), "Establishing an information security culture in organizations: an outcomes based education approach", M dissertation, Nelson Mandela Metropolitan University, Port Elizabeth.
- van Niekerk, J.F. and von Solms, R. (2004), "Corporate information security education: is outcomes based education the solution?", paper presented at the 10th IFIP WG11.1 Annual Working Conference on Information Security Management, World Computer Congress (WCC), Toulouse.
- von Solms, R. (1998), "Information security management (3): the code of practice for information security management (BS7799)", *Information Management & Computer Security*, Vol. 6 No. 5, pp. 224-5.

About the authors

Hennie Kruger is an Associate Professor in the School of Computer, Statistical and Mathematical Sciences at the North-West University (Potchefstroom Campus) in South Africa. He previously worked for Anglo American Corporation as a senior Computer Auditor and has more than ten years experience in Information Risk Management. He has a PhD in Computer Science, an MCom in Information Systems and an MSc in Mathematical Statistics. His current interests include decision modeling, security awareness and the use of linear programming models. Hennie Kruger is the corresponding author and can be contacted at: hennie.kruger@nwu.ac.za

Lynette Drevin is a Lecturer in Computer Science and Information Systems at the North-West University (Potchefstroom Campus) in South Africa. She joined the university in 1985 as a staff member. She has a MSc in Computer Science and Information Systems. Her current interests include security awareness and education and project failures and the use of narratives.

Tjaart Steyn is a Professor in Computer Science in the School of Computer, Statistical and Mathematical Sciences at the North-West University (Potchefstroom Campus) in South Africa. He joined the university in 1974 as a staff member of the then IT department and moved to Computer Science in 1980. He has a PhD in Computer Science. His current interests include cutting stock optimization, security awareness and databases.

To purchase reprints of this article please e-mail: reprints@emeraldinsight.com
Or visit our web site for further details: www.emeraldinsight.com/reprints