A Linguistic Approach to Information Security Awareness Education in a Healthcare Environment

Lynette Drevin, Hennie Kruger, Anna-Marie Bell, Tjaart Steyn

A linguistic approach to information security awareness education in a healthcare environment

L Drevin, H Kruger, A Bell and T Steyn

North-West University, Computer Science & Information Systems, Potchefstroom, South Africa

Lynette.Drevin@nwu.ac.za, Hennie.Kruger@nwu.ac.za, AnnaMarie.Bell@nwu.ac.za, Tjaart.Steyn@nwu.ac.za

Abstract: It is widely accepted that healthcare information security is extremely important and that security breaches will have serious consequences in many areas. Despite controls, such as legal frameworks, as well as ongoing research projects into healthcare information security and privacy, there is still an alarming number of healthcare information security breaches reported annually. In this paper, a linguistic approach, utilizing a vocabulary test, is proposed as a tool to determine security awareness levels of healthcare workers and to assist in educating them in security awareness aspects. A vocabulary-measuring instrument was developed and distributed to healthcare workers in a large South African hospital group. Results indicated that information security awareness levels are generally acceptable, but that potential problem areas exist between certain language groups, as well as between different business functions (departments). The study also shows that the proposed approach may offer significant advantages in information security awareness campaigns.

Keywords: Information security awareness, vocabulary test, healthcare, linguistics, behavior, knowledge, education.

1 Introduction

Healthcare information systems (HIS) play a critical role within healthcare establishments and have become an integral part of all aspects of modern healthcare. These types of systems operate in a connected and networked environment and receive, store, send and process data that are generally deemed to be of an extremely sensitive nature. A healthcare information system is primarily centered towards the patient [1] and examples of electronic resources in HIS may include digital patient records, diagnostic and treatment data, service provider information, financial information (e.g. medical schemes), etc. Security breaches of healthcare information may have serious consequences in many areas. Personal health information is only one of the areas and patients may be exposed to economic threats, mental anguish and possible social stigma should information, such as medical history, test and laboratory

results or insurance information leak out [2]. All these point towards an increasing need for security and privacy of electronic health records.

The protection of sensitive electronic health records is not only required as a best practice, but should also adhere to legal liabilities. Maseti [3] pointed out that there are well over 30 countries that have enacted information protection statutes at national or federal level. With regard to the protection of health information, the Health Insurance Portability and Accountability Act of 1996 (HIPAA) in the United States of America is probably of the most well-known legislation [4]. The act ensures that consistent standards are maintained with regard to the handling and privacy of medical information records [5]. In South Africa, where this study was performed, a number of acts are in place to address the security and privacy of information. The Promotion of Access to Information Act No 2 of 2000 (PAIA) provides for giving access to a person to his or her information to ensure correctness and accuracy; The Electronic Communications and Transaction Act No 25 of 2002 deals directly with the protection of personal information; and The Protection of Personal Information Act No 4 of 2003 (POPI) regulates how personal information should be handled, stored and secured [6]. This act applies to anyone in South Africa who processes personal information [3].

The importance of healthcare information and the protection of such information are further emphasized by the large number of research projects and research papers on the topic. Examples of related work that investigate various aspects of healthcare information security include the following. Medlin and Cazier [7] explored the use of social engineering techniques and associated password practices of healthcare workers; Van Deursen et al. [8] applied a mixed methods approach to identify information security risks within healthcare; Appari and Johnson [9] surveyed the literature and provided a holistic view of research in information security and privacy in healthcare. They also suggest new areas that may be of interest in healthcare information security; and, Fernando and Dawson [10] performed a case study to suggest a health information system security threat life cycle. An informative overview of the security and privacy of electronic health records, based on a systematic literature review, is presented in Fernandez-Aleman et al. [11]. Earlier studies can also be found in [12], [13] and [14].

Despite the acknowledgement that healthcare information security is extremely important, as well as the existence of a legal framework to protect healthcare information, and the ongoing research into healthcare information security and privacy, there is still an alarming number of security breaches, which are regularly reported in various reports. The 2015 breach report on protected health information [15] reports that an astonishing more than 113 million patient health records were breached in 2015. This represents an 897% increase in records breached from 2014 to 2015. It is also significant to note that 98.1% of the records breached in 2015 was the result of hacking attacks or other information technology incidents. These statistics are consistent with those reported in other reports [16], [17].

The above introductory comments clearly indicate an ongoing need to educate healthcare workers and to make them aware of information security threats.

Information security awareness plays a significant role in combatting undesirable information technology incidents. Ogutcu et al. [18] pointed out that undesirable behavior is closely related to an IS user's level of information security awareness and that every person is obliged to be aware of information security threats in order to protect information. To design and implement effective information security awareness programs requires the identification of appropriate security topics to be included. Katsikas [19], for example, suggested a methodology to determine information security training needs specifically aimed at different personnel groupings within healthcare establishments. In this paper a linguistic approach, utilizing a vocabulary test, is described to identify areas of concern in information security awareness levels of healthcare workers. The study, which was performed within a large hospital group in South Africa, is based on an earlier exploratory study where the use of a vocabulary test in the context of security awareness was tested with students at a university [20].

The remainder of the paper is organized as follows. In Section 2, a brief theoretical background of the linguistic approach followed is presented. In Section 3, the methodology used in the study is described and in Section 4, the results and a discussion of the results are presented. The paper is then concluded with some final remarks in Section 5.

2 Theoretical context

Linguistics is the scientific study of human language and linguistic knowledge is called a grammar [21]. According to Fromkin [21], a grammar includes various concepts, such as a lexicon (vocabulary), morphology (structure of words), syntax (structure of phrases and sentences), semantics (meaning of words) and phonology (sounds). This study proposes a linguistic approach to assess information security awareness and focuses in particular on the lexicon or vocabulary component, as well as, to a certain extent, the semantics of linguistics.

The proposed linguistic framework may be seen as an expansion of the work reported in Kruger et al. [20]. In this initial exploratory project, the feasibility of a vocabulary test as a tool to assess security awareness was investigated. Empirical tests were performed on students at a tertiary institution and results proved to be useful. Utilizing vocabulary tests in information security is not generally practiced and stems from the success of other educational studies (specifically mathematics) where learners' language proficiency in a subject, e.g. mathematics, was measured using vocabulary tests [22], [23].

A detailed discussion of language development, vocabularies and the associated processes to transform linguistic inputs into beliefs and actions is beyond the scope of this paper. Such a discussion will have to include, amongst others, aspects such as cognitive psychology. Details on these topics can be found in [24] and, to a lesser extent, in [20]. For the purposes of this paper, it will suffice to state that cognitive principles (learn, know, understand, process and recall of information) are important

in linguistic tests. Van der Walt [25] argues that three key cognitive skills are necessary for successful learning. The three skills are knowledge of facts, processes and concepts, the ability to apply the knowledge, concepts and processes, and the ability to reason. These three cognitive skills are explained in the context of information security in Table 1. Table 1, which was originally adapted from [25], is quoted directly from [20], as both the original study and this current study are based on the same cognitive skills.

Table 1. Cognitive skills [20]

Cognitive category	Cognitive action
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
Explanation: 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 enables people to see the relationship among the different elements of information security and helps 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
Explanation: 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
Explanation: Reasoning in information security requires logical and systematically, including intuitive and inductive, thinking processes. People should be able to implement expertise in different contexts.	

Based on the above brief comments and the discussion in Section 1, it was decided to perform a comprehensive vocabulary test exercise on the staff of a large hospital group in an effort to determine information security awareness levels. The study was motivated by the importance of medical information and the associated risks and threats that occur in healthcare environments (see Section 1). Choosing a linguistic approach is justified by the success of the initial study to use a vocabulary test in the context of information security awareness [20]. This approach is of particular interest to South Africa where the study was performed. South Africa is a multilingual (11 official languages) country where information is often presented in only one language and a real danger of misunderstanding or misinterpreting information and messages exists. Furthermore, it is assumed that information system users may be more susceptible to security breaches if they do not have a basic understanding and comprehension of information security concepts and terminology. An additional advantage of using a vocabulary test is the opportunity it creates to identify suitable topics for inclusion in an information security awareness program as well as the

identification of appropriate techniques to conduct awareness campaigns. It should, however, be noted that a linguistic approach may produce unreliable results in some instances, e.g. normal ethical behavior may prevent users from revealing confidential information without knowing what the concept "social engineering" entails. To address this type of problem the measuring instrument also provides for scenario-type questions to test a user's behavior. Details on this and the general methodology followed are highlighted in the next section.

3 Methodological approach

The general methodology followed in this study was based on a vocabulary-measuring instrument. A questionnaire, consisting of two sections, was developed to assess the security awareness levels of information technology users in a large hospital group. The aim of the first section was to test specific information technology-related knowledge, while the second section focused on the evaluation of expected information security behavior.

To compile the knowledge section of the questionnaire, an extensive Internet search was conducted to identify a list of appropriate information security-related words or concepts. This resulted in a list of 45 concepts that were categorized in four main categories, namely social engineering, viruses, spam, and unauthorized access. Following a rigorous review of these concepts and performing small pilot studies, it was decided to use 20 of the 45 words in the final measuring instrument. An example of a vocabulary-type question to test a respondent's information security knowledge is given below.

Example 1. Vocabulary-type question

A computer virus is a:

(a) Computer program that is designed to replicate itself by copying itself into other programs in a computer. It may be benign or have a negative effect, such as causing other programs to operate incorrectly or corrupting a computer's memory

(b) Term used to describe any computer program that functions incorrectly

(c) Computer program that will influence your computer's performance and which can be bought at any reputable computer dealer

(d) Group of unsolicited bulk email messages creating havoc

(e) I do not know

The second section of the questionnaire contained 10 scenario-type questions. The purpose of these questions was to establish whether users really understand a specific information security concept or whether behavior was merely a result of normal underlying ethical principles. The scenario-type questions were based on appropriate

examples in the literature [20], [26], as well as information obtained during pilot studies. Example 2 indicates a scenario-type question.

Example 2. Scenario-type question

When you receive a recorded phone call requesting you to call a toll-free telephone number that purports to be that of a well-known financial institution and you are asked to punch in any personal information, what would you do?
(a) Ignore the request
(b) End the telephone call
(c) Phone the toll-free number and supply the information
(d) None of the above
(e) I do not know

In addition to the Internet searches and literature reviews, the measuring instrument was also subjected to a number of pilot studies. The initial questionnaire was handed to eight staff members of an IT department at a university for comments. Certain changes were suggested and the improved questionnaire was then tested with another 31 respondents. This sample of respondents represented a mix of senior students and lecturers from the same university. The questionnaire was again adjusted according to the feedback of the respondents. A final pilot study was then conducted with the hospital group where the actual study was performed. A total of 46 staff members of the hospital group participated in this pilot study. During the process of constructing and refining the measuring instrument, basic statistical tests were also performed to ensure validity and reliability of the questions included in the final instrument. A limited number of these statistical test results will be highlighted in the next section where the results are presented.

The final questionnaire was made available electronically for two months to information technology users within a large South African hospital group. The hospital group consists of 55 hospitals and 48 retail pharmacies with over 29 000 employees. A total of 3 577 staff members across different departments (e.g. Finance, Marketing, Information Technology, Production, Purchasing, Customer Services, etc.) received the questionnaire by means of their management structure. From this population, 1 039 respondents completed the questionnaire. The questionnaire is not included due to confidentiality reasons. A summary of some of the results are presented in the next section.

4 Results and discussion

Due to space considerations not all results can be reported here and it was decided to present two sets of results that are of particular interest. The first set of results pertains to the different language groups of the hospital's workforce, while the second set highlights some of the results linked to the different business activities.

The main demographic details of the 1 039 respondents are shown in Table 2.

Table 2. Demographic details of respondents

Gender			**Main function or activity**	
Female	**77.3%**		Management	**33.5%**
Male	22.7%		Other	27.1%
			Clerical	24.1%
Age			Patient Care	11.3%
Younger than 20	0.3%		Auditing	3.9%
21 - 30	21.4%		**Home language**	
31 - 40	**33.9%**		Afrikaans	**44.2%**
41 - 50	26.8%		English	37.9%
Older than 50	17.5%		Other	17.9%

Basic statistical tests were performed on the complete data set to ensure that valid and reliable results have been obtained. Reliability (accuracy of the measuring instrument) was assessed by calculating the well-known Cronbach Alpha coefficient [27]. The resulting coefficient was 0.71 and based on this, the results were accepted as reasonable. To test for statistical significance (and significant differences between groups of people, e.g. first language groups or business activity groups), it was decided to make use of the Cohen's d-value – also known as an effect size [28]. The d-value is defined as a standardized difference between the means of two populations (groups in this study) and gives the importance of the effect in practice. The standard interpretation of the d-value is given by Cohen [28] as 0.8 = large effect; 0.5 = medium effect; and 0.2 = small effect. This statistic was used to test whether there are significant differences between the information security awareness knowledge of the different language and business function groups. Table 3 presents the d-values of information security awareness knowledge per language group.

Table 3. Effect size (d-values) – Knowledge per language group

Effect size (d-values)		
Afrikaans with English	Afrikaans with Other	English with Other
0.13	0.49	0.61

It should be noted that there are 11 official languages in South Africa. For reporting purposes and due to the number of respondents (per language group), it was decided to group all languages, other than Afrikaans and English, together as "Other". The reported d-values indicated that there is a medium effect (d=0.49), or medium difference, between the information security knowledge of Afrikaans-speaking health-care workers and those who speak one of the other official languages. There is, however, a more significant difference (d=0.61) between the knowledge of English-speaking workers and those classified as other. This finding is of particular interest as the measuring instrument is focused on linguistic aspects. The differences in

knowledge from the different language groups may indicate that language is probably a barrier in security awareness programs where information security concepts are explained. To further illustrate this finding, consider the graph in Figure 1. Figure 1 graphically displays the result of a typical knowledge question where respondents had to explain the meaning of the term "identity theft".

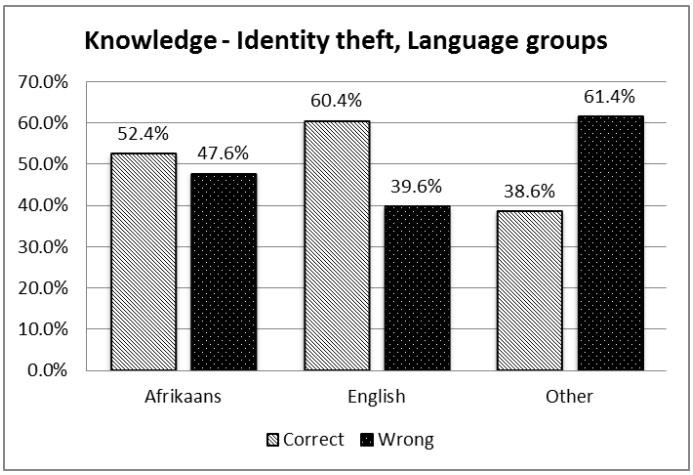

Figure 1. Vocabulary-type question

It follows from Figure 1 that 52.4% of Afrikaans-speaking workers responded in the correct way, 60.4% of English-speaking workers were correct and only 38.6% of the other language groups answered correctly. Overall just over half (53.2%) of respondents knows what "identity theft" is.

In terms of different business functions there were only medium differences recorded in knowledge. Medium effect sizes were noted between the business functions Clerical and Management (d=0.37); Auditing and Management (d=0.40); and Patient Care and Management (d=0.41). What is of significance here is that Management seems to be the common denominator when it comes to knowledge differences in information security concepts. Management have consistently scored higher than the other groups. This may be an indication that information security awareness material should be different for different business function groups. As an additional example to this finding, consider the results of a scenario-type (behavior) question per business function as depicted in Figure 2.

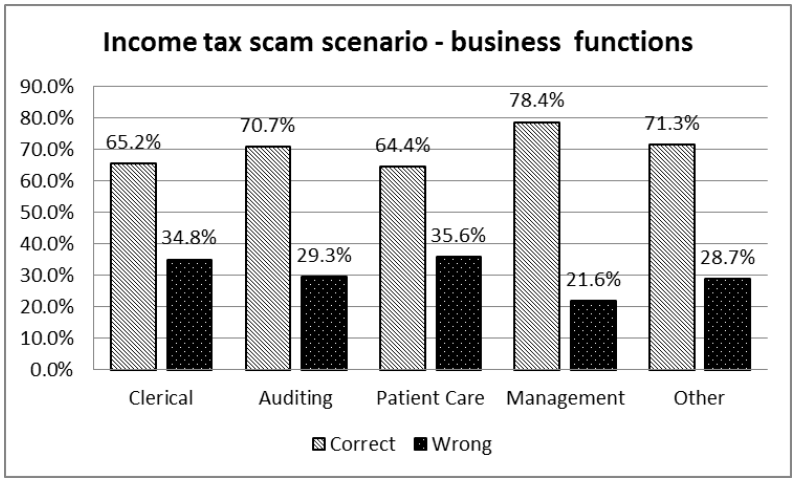

Figure 2. Scenario-type question

Respondents had to react to an Income Tax scam scenario. Although 71% of respondents indicated that they would behave in an acceptable and secure manner, it is significant that there are notable differences between business functions, e.g. Patient Care (64.4%) versus Management (78.4%).

The results presented in this section represent only a small extract of the total results of the study. There were also other examples where differences and poor knowledge were recorded. Conversely, there were also a number of results that indicated that healthcare workers do have good knowledge of specific information security concepts and appropriate behavior. The aim was to show that a linguistic approach does offer advantages and that it is possible to use such an approach to collect important management information pertaining to information security awareness. This study indicates that specific terms and concepts, as well as behavior need to be clarified and should focus on an information security awareness program for at least those healthcare workers who participated in the study. Furthermore, home language and type of work (business function) play a definite role. Certain groups are more prone to risky behavior and certain groups have less information security knowledge. This information is helpful to focus on specific information security topics and concepts, as well as on security-related behavior.

5 Conclusions

Cyberattacks on healthcare organizations are now a fact of life [17]. Healthcare workers depend more and more on healthcare information systems and information security and privacy have become issues of growing importance. This clearly indicates an ongoing need for information security education, and particularly for ensuring high levels of information security awareness among healthcare workers.

This paper presents a linguistic approach, utilizing a vocabulary test, to evaluate healthcare workers' information security levels and to identify possible areas of concern. A vocabulary-measuring instrument was developed and made available to a large South African hospital group with over 29 000 employees. Although positive results were recorded in many instances, the study has shown that potential problem areas exist in terms of different language groups and different business functions. Based on the results it was also concluded that the use of a linguistic approach does offer advantages in information security awareness programs.

Future work may include updating the questionnaire with recent and new threats as the security landscape has a dynamic nature. This survey can also be conducted in other industries for comparative studies.

References

1. Haux, R.: Health information systems – past, present, future. International Journal of Medical Informatics, 75:268-281 (2006)
2. Appari, A. and Johnson, M.E.: Information security and privacy in healthcare: Current State of Research. Center for Digital Strategies Tuck School of Business Dartmouth College (2008)
3. Maseti, O.: A model for role-based security education, training and awareness in the South African healthcare environment. Unpublished M dissertation. Nelson Mandela Metropolitan University (2008)
4. Win, K.T.: A review of security electronic health records. Health Information Management, 34(1):13-18 (2005)
5. Meingast, M., Roosta, T. and Sastry, S.: Security and privacy issues with health care information technology. Proceedings of the 28th IEEE EMBS Annula International Conference. New York City, USA (2006)
6. South Africa.: Protection of Personal Information Act No 4 of 2013. Government gazette, 581:37067, 26 November 2013 (2013)
7. Medlin, B.D. and Cazier, J.A.: Social engineering techniques and password security: two issues relevant in the case of health care workers. IGI Global, 3(2):58-70.
8. Van Deursen, N., Buchanan, W.J. and Duff, A. 2013. Monitoring information security risks within health care. Computers & Security, 37:31-45 (2014)
9. Appari, A. and Johnson, M.E.: Information security and privacy in healthcare: current state of research. International journal of Internet and enterprise management, 6(4):279-314 (2010)
10. Fernando, J.L. and Dawson, L.L.: The health information system security threat lifecycle: an informatics theory. International Journal of Medical Informatics, 78:815-826 (2009)
11. Fernandez-Aleman, J.L., Senor, I.C., Lozoya, P.A.O. and Toval, A.: Security and privacy in electronic health records: A systematic literature review. Journal of Biomedical Informatics, 46:541-562 (2013)
12. Cavalli, E., Mattasoglio, A., Pinciroli, F. and Spaggiari, P.: Information security concepts and practices: the case of a provincial multi-specialty hospital. International Journal of Medical Informatics, 73:297-303 (2004)
13. Janczewski, L. and Shi, F.X.: Development of information security baselines for healthcare information systems in New Zealand. Computers & Security, 21:172-192 (2002)
14. Smith, E. and Eloff, J.H.P.: Security in health-care information systems – current trends. International Journal of Medical Informatics, 54:39-54 (1999)

15. Redspin. 2016. Breach report 2015: Protected Health Information (PHI). Redspin.
16. Munro, D. 2015. Data breaches in healthcare totaled over 112 million records in 2015. Forbes, http://www.forbes.com/sites/danmunro/2015/12/31/data-breaches-in-healthcare-total-over-112-million-records-in-2015/#1996c0557fd5
17. HIPAA. 2016. Major 2016 healthcare data breaches: midyear summary. HIPAA Journal, http://www.hipaajournal.com/major-2016-healthcare-data-breaches-mid-year-summary-3499/
18. Ogutcu, G., Testik, O.M. and Chouseinoglou, O.: Analysis of personal information security behavior and awareness. Computers & Security, 56:83-93 (2016)
19. Katsikas, S.K.: Health care management and information systems security: awareness, training or education? International Journal of Medical Informatics, 60:129-135 (2000)
20. Kruger, H.A., Drevin, L. and Steyn, T.: A vocabulary test to assess information security awareness. Information Management & Computer Security, 18(5):316-327 (2010)
21. Fromkin, V.A. (ed.): Linguistics: An introduction to linguistic theory. Fromkin, V.A. Wiley-Blackwell (2001)
22. Jansen van Vuuren, N.: 'n Ondersoek na die gebruik van wiskunde-woordeskat en metakognitiewe strategieë tydens probleemoplossing by Graad 7-leerders.Unpublished M dissertation, North-West University, Potchefstroom (2014)
23. Van der Walt, M.S., Maree, K. and Ellis, S.: A mathematics vocabulary questionnaire for use in the intermediate phase. South African Journal of Education, 28:489-504 (2008)
24. Robinson-Riegler, G. and Robinson-Riegler, B.: Cognitive psychology. Applying the science of mind. Boston MA: Pearson (2008)
25. Van der Walt, M.S.: Aanpassing van die studie oriëntasievraelys in Wiskunde vir gebruik in die intermediêre fase. Unpublished PhD dissertation, North-West University, Potchefstroom (2008)
26. Furnell, S.M., Bryant, P. and Phippen, A.D.: Assessing the security perceptions of personal Internet users. Computers & Security, 26(5): 410-417 (2007)
27. Kerlinger, F.N.: Foundations of Behavioral Research. Third edition. Japan: CBS Publishing (1986)
28. Cohen, J.: Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates (1988)