

# **Evaluating Readability as a Factor in Information Security Policies**

**By**

**Yazeed Ahmed Alkhurayyif**

**Supervisor: Dr. George Weir**

Thesis submitted for the degree of Doctor of Philosophy

School of Computer and information

Strathclyde University

2019

## List of Publications

### Posters

1. Alkhurayyif, Y. and Weir, G. R. S. (2015) ‘How can we gauge the effectiveness of information security policies?’, in *SICSA\_2015*. University of Glasgow.
2. Alkhurayyif, Y. and Weir, G. R. S. (2016) ‘Understanding what makes security policies effective’, in *IAARHIES\_2016*. London.
3. Alkhurayyif, Y. and Weir, G. R. S. (2016) ‘Document analysis as a basis for information Security policy assessment’, in *the Forensic science in defence Symposium*. University of Cranfield.
4. Alkhurayyif, Y. and Weir, G. R. S. (2017) ‘How can we measure the likely effectiveness of information security policies?’, in *SIPR/Police Scotland postgraduate symposium*. University of Edinburgh.
5. Alkhurayyif, Y. and Weir, G. R. S. (2018) ‘Estimating readability of ISPs with traditional readability metrics and recent statistical approaches’, in *SICSA\_2018*. University of Robert Gordon University.
6. Alkhurayyif, Y. and Weir, G. R. S. (2018) ‘Can readability metrics help to improve security policies?’, in *5<sup>th</sup> Strathclyde workshop on cybersecurity*. University of Strathclyde.

### Papers

7. Alkhurayyif, Y. and Weir, G. R. S. (2017) ‘Readability as a basis for information security policy assessment’, in *2017 Seventh international conference on emerging security technologies (EST)*. IEEE, pp. 114–121. doi: 10.1109/EST.2017.8090409. Available at: <http://ieeexplore.ieee.org/document/8090409/>
8. Alkhurayyif, Y. and Weir, G. R. S. (2017) ‘Evaluating readability as a factor in information security policies’, in *International journal of trend in research*

*and development*, (December), pp. 54–64. Available at:  
<http://www.ijtrd.com/papers/IJTRD14635.pdf>.

9. Alkhurayyif, Y. and Weir, G. R. S. (2018) ‘Using sequential exploratory mixed methods design to explore readability of ISPs’, in *2018 International conference on computing, electronics & communications engineering (iCCECE)*. IEEE, pp. 123-127. Available at:  
<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8659025>

## **Awards**

1. The poster entitled ‘Understanding what makes security policies effective’ has been awarded as the best poster in the IAARHIES 40th International Conference in London.
2. I have been awarded as a Saudi distinguished student from his highness prince Mohammed bin Nawaf, the ambassador of the Saudi Arabia Kingdom to the United Kingdom, in 2017.
3. The paper entitled ‘Evaluating readability as a factor in information security policies’ has been awarded as the best paper authored by a research student in 2017 (Euan Minto Prize) on computer and information science department, University of Strathclyde.

## **Abstract**

Policies should be treated as rules or principles that individuals can readily comprehend and follow as a pre-requisite to any organisational requirement to obey and enact regulations. This dissertation attempts to highlight one of the important factors to consider before issuing any policy that staff members are required to follow. Presently, there is no ready mechanism for estimating the likely efficacy of such policies across an organisation. One factor that has a plausible impact upon the comprehensibility of policies is their readability. Researchers have designed a number of software readability metrics that evaluate how difficult a passage is to comprehend; yet, little is known about the impact of readability on the interpretation of information security policies and whether analysis of readability may prove to be a useful insight.

This thesis describes the first study to investigate the feasibility of applying readability metrics as an indicator of policy comprehensibility through a mixed methods approach, with the formulation and implementation of a seven phase sequential exploratory fully mixed methods design. Each one was established in light of the outcomes from the previous phase. The methodological approach of this research study is one of the distinguishing characteristics reported in the thesis, which was as follows:

- eight policies were selected (from a combination of academia and industry sector institutes);
- specialists were requested their insights on key policy elements;
- focus group interviews were conducted;
- comprehension tests were developed (Cloze tests);
- a pilot study of comprehension tests was organised (preceded by a small-scale test);
- a main study of comprehension tests was performed with 600 participants and reduce that for validation to 396;
- a comparison was made of comprehension results against readability metrics.

The results reveal that the traditional readability metrics are ineffective in predicting human estimation. Nevertheless, readability, as measured using a bespoke readability metric, may yield useful insight upon the likely difficulty that end-users may face in comprehending a written text. Thereby, our study aims to provide an effective approach to enhancing the comprehensibility of information security policies and afford a facility for future research in this area.

The research contributes to our understanding of readability in general and offering an optimal technique to measure the readability in particular. We recommend immediate corrective actions to enhance the ease of comprehension for information security policies. In part, this may reduce instances where users avoid fully reading the information security policies, and may also increase the likelihood of user compliance. We suggest that the application of appropriately selected readability assessment may assist policy makers to test their draft policies for ease of comprehension before policy release. Indeed, there may be grounds for a readability compliance test that future information security policies must satisfy.

**Keywords:** *Readability; Readability formula, Readability metric, Information security policy, Information security, Comprehension test*

## **Attestation**

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyright Acts as qualified by University of Strathclyde Regulation 3.50. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

Signed:

Date:

## **Acknowledgements**

First and above all, I would like to give honour and praise to the Almighty Allah for giving me strength and his guidance throughout my thesis process.

I would like to express my deep appreciation to my father, who, although no longer with us, played a major role in inspiring me in doing the PhD degree. His wills are the constant motivation to continue my doctoral study. I dedicate this thesis to my only late brother, Abdulmohsen, for his emotional support.

My sincere gratitude and appreciation to my supervisors, Dr George Weir and Dr John Wilson for their support, guidance and leadership during this journey. Their immense patience, charm, insights, and professionalism helped me to conquer several challenges over the past few years.

My exceptional appreciation to my mother for her enthusiasm, support and prayers. My heartfelt thanks and deepest love to my dear wife for her continued sacrifices, care, and encouragement. Without them, everything would have been much harder. However, the time has come to celebrate and pursue our life together with my son Rayyan.

I am grateful to my relatives and friends for their continuous support. My sincerest thanks are due to participating individuals for their time, comments and participation in this research.

My last words are for my PhD financial sponsor, as this work has been made possible by the Shaqra University scholarship, for which I am very grateful.

# Table of Contents

List of Publications .....	II
Posters .....	II
Papers .....	II
Awards .....	III
Abstract .....	IV
Attestation .....	VI
Acknowledgements .....	VII
Table of Contents .....	VIII
List of Tables.....	XIV
List of Figures .....	XV
<b>CHAPTER ONE: INTRODUCTION .....</b>	<b>1</b>
1.1.    Introduction.....	1
1.2.    Outline of the Thesis .....	3
<b>CHAPTER TWO: BACKGROUND AND RELATED RESEARCH.....</b>	<b>4</b>
2.1.    Introduction.....	4
2.2.    The Need for Security .....	4
2.3.    Information Security Background.....	6
2.4.    Information Security .....	7
2.5.    Information Security Principles .....	8
2.6.    Security Approaches .....	9
2.6.1.    Technical Approaches .....	10
2.6.2.    Non-technical Approaches.....	14
2.7.    Rewards or Sanctions Approach .....	16



2.8.	Information Security: Threats, Vulnerabilities and Countermeasures.	18
2.9.	Information Security and the Institution .....	20
2.10.	Information Security in Higher Education.....	21
2.11.	Raising Awareness.....	22
2.12.	The Importance of Information Security Awareness Programmes.....	25
2.13.	Obstacles to Establishing Information Security Awareness Programme	26
2.14.	The Effectiveness of Information Security Awareness Programme ....	27
2.15.	Programme Drivers .....	28
2.16.	What is an Information Security Policy? .....	29
2.17.	The Role and Importance of the Information Security Policy .....	30
2.18.	Compliance with the Information Security Policy.....	31
2.18.1	Threats to Compliance.....	32
2.19.	How to Make Successful Information Security Policies.....	34
2.20.	Evaluating the Success and Effectiveness of Information Security Policies	35
2.21.	Summary .....	37
<b>CHAPTER THREE: READABILITY.....</b>		<b>39</b>
3.1.	Introduction.....	39
3.2.	Defining Readability .....	39
3.3.	How to Evaluate Readability Measurement?.....	41
3.3.1.	Standard Approach to Readability Measurements.....	41
3.3.2.	Another Approach to Readability.....	51
3.4.	Comparison of Approaches to Readability .....	53
3.5.	Summary .....	58
<b>CHAPTER FOUR: RESEARCH METHODOLOGY .....</b>		<b>59</b>

4.1.	Introduction .....	59
4.2.	Understanding Methodology .....	59
4.3.	Research Approaches .....	60
4.3.1.	Philosophical Worldviews .....	61
4.3.2.	Research Designs .....	65
4.3.3.	Research Methods .....	68
4.4.	Choice of Research Approaches .....	69
4.4.1.	Qualitative Methods .....	70
4.4.2.	Quantitative Methods .....	72
4.5.	Research Design (Exploratory Sequential) .....	73
4.5.1	Why the Exploratory Sequential Approach is Selected .....	73
4.6.	Ethical Assurance .....	76
4.7.	Summary .....	77
<b>CHAPTER FIVE: EXPERIMENTAL DESIGN .....</b>		<b>78</b>
5.1.	Introduction .....	78
5.2.	The Research Methodology Steps .....	78
5.2.1	Phase1 .....	79
5.2.2	Phase2 .....	79
5.2.3	Phase3 .....	79
5.2.4	Phase4 .....	80
5.2.5	Phase5 .....	81
5.2.6	Phase6 .....	81
5.2.7	Phase7 .....	82
5.3.	Information Security Policies Documents .....	82
5.3.1	Core Process .....	82
5.4.	Prior to Experts' Insights Experiment .....	84

5.4.1	Preparing Draft Instrument.....	84
5.4.2	Procedure.....	84
5.4.3	Participants.....	84
5.4.4	Results.....	85
5.4.5	Summary.....	85
5.5.	Experts' Insights Experiment.....	85
5.5.1	Procedures.....	86
5.5.2	Participants.....	87
5.5.3	Conclusion.....	88
5.6.	Focus Group Methodology.....	88
5.6.1	Focus Discussion Group Report.....	89
5.6.2	Procedure Preparation.....	90
5.6.3	Methodology.....	91
5.6.4	Results.....	93
5.6.5	Conclusion.....	94
5.7.	Pilot Study of Cloze Tests.....	95
5.7.1	Aims.....	95
5.7.2	Participants.....	96
5.7.3	Materials.....	96
5.7.4	Procedures.....	97
5.7.5	The Pilot Study Results Implication for the Main Study.....	98
5.7.6	Conclusion.....	100
5.8.	Second Pilot Study.....	101
5.8.1	Aims.....	101
5.8.2	Why LearnClick Instrument was Chosen.....	101
5.8.3	Participants.....	103

5.8.4	Materials .....	103
5.8.5	Procedures .....	104
5.8.6	Data Collection.....	104
5.8.7	Data Analysis.....	105
5.8.8	Implications for the Main Study.....	105
5.8.9	Results .....	108
5.8.10	Conclusion .....	110
5.9.	Main Study .....	111
5.9.1	Aims .....	111
5.9.2	Questions .....	111
5.9.3	Sub-questions .....	112
5.9.4	Demographic Questions .....	112
5.9.5	Sample Size .....	113
5.9.6	Participants .....	114
5.9.7	Materials .....	115
5.9.8	Procedures .....	118
5.9.9	Measures.....	121
5.9.10	Data Collection .....	122
5.10.	Summary .....	124
<b>CHAPTER SIX: RESULTS AND DATA ANALYSIS.....</b>		<b>126</b>
6.1.	Introduction.....	126
6.2.	Descriptive Statistics.....	126
6.3.	Results.....	127
6.3.1	Demographics.....	127
6.3.2	Participants' Results VS Traditional Readability Index's Outcomes	141

6.3.3	Participants' Cloze Test Results VS Strathclyde Readability Measure1	
	Outcomes	143
6.3.4	Key Findings of Demographic Variables.....	145
6.3.5	The Impact of Demographic Variables on Reading Comprehension	
		146
6.4.	Summary .....	150
<b>CHAPTER SEVEN: DISCUSSION AND CONCLUSION .....</b>		<b>152</b>
7.1.	Introduction.....	152
7.2.	Discussion .....	152
7.1.1	Summary of the Main Study's Findings.....	152
7.2.1	An Indication of the Importance of the Findings .....	156
7.3.	Contribution and Implications .....	157
7.3.1	Theoretical and Practical Contributions .....	157
7.3.2	Practical Contributions .....	158
7.4.	Limitations and Suggestions for Future Research .....	159
7.5.	Recommendations .....	160
7.6.	Summary .....	160
	REFERENCES.....	161
	APPENDIX (A) .....	186
	APPENDIX (B).....	187
	APPENDIX (C).....	194

## List of Tables

Table 2. 1: The advantages and disadvantages of authenticators.....	11
Table 3. 1: Factors that can affect readability .....	40
Table 3. 2: Predictor variables used in the eight readability metrics .....	43
Table 3. 3: Comparison between readability automated tools .....	58
Table 4. 1: Four philosophical worldviews.....	65
Table 4. 2: Research designs' types .....	68
Table 4. 3: Research methods .....	69
Table 5. 1: Comparison of human results & SRM 1 results .....	110
Table 5. 2: The main study's demographic questions.....	120
Table 6. 1: The demographic information.....	133
Table 6. 2: Response by nationality .....	135
Table 6. 3: The mean of the survey elements.....	138
Table 6. 4: Overview of the examined policies results .....	140
Table 6. 5: Comparison of human results VS traditional formulae results.....	143
Table 6. 6: Comparison of human results VS SRM 1 results .....	144
Table 6. 7: Analysis results on demographic questions .....	150
Table A3. 1: Estimate of readability on the FRE and SRM scale.....	186
Table A5. 1: The second pilot study's demographic information.....	190
Table A5. 2: The second pilot study's participants results .....	191
Table A5. 3: The second pilot study's descriptive statistics .....	191
Table A6. 1: Comparison of human results VS SRM 1 results .....	194

## List of Figures

Figure 3. 1: Cloze test's types example .....	52
Figure 4. 1: A framework for research extracted from (Creswell and Creswell, 2018, p. 5) .....	61
Figure 4. 2: Three mixed methods designs, extracted from (Creswell and Creswell, 2018, p. 218) .....	67
Figure 4. 3: The exploratory sequential design for the thesis .....	75
Figure 4. 4: Our research framework .....	75
Figure 5. 1: The research methodology steps.....	78
Figure 5. 2: Experts insight experiment process .....	87
Figure 5. 3: Focus group process .....	93
Figure 5. 4: Example of demographic question types.....	122
Figure 5. 5: The main study process .....	124
Figure 6. 1: Response by gender .....	128
Figure 6. 2: Response by language .....	128
Figure 6. 3: Response by number of years for studying the English language.....	129
Figure 6. 4: Response by age .....	129
Figure 6. 5: Response by qualification level.....	130
Figure 6. 6: Response by computer experience level.....	131
Figure 6. 7: Response by study subject.....	131

# **CHAPTER ONE: INTRODUCTION**

## **1.1. Introduction**

Presently, there is no ready mechanism for estimating the likely efficacy of information security policies (ISPs) across an organisation. One factor that has a plausible impact upon the comprehensibility of policies is their readability. There are many readability formulae that aim to assess how difficult a text document is to comprehend; so far, little is known about the effect of readability on the interpretation of ISPs and whether examination of readability may offer a useful insight.

This research is the first study to examine the effectiveness of employing readability metrics as an indicator of policy comprehensibility through a mixed methods approach, with the formulation and implementation of a seven-phase sequential exploratory fully mixed methods design. Most previous studies were focused on comparing readability formulae without being tested empirically by a software aspect and human aspect. This study identifies that the traditional readability metrics are ineffective in predicting the human estimation. This thesis contributes to current knowledge regarding the level of analysis. This research involved a variety of participant's groups (ISP experts, professional computer users, researchers, local students and international students in the United Kingdom). The original data in this study were collected from nearly 450 participants by a variety of instruments (pilot studies, main study, focus groups, and expert insights). This integration presents an indication of the complexity that researchers face if they are to succeed in covering all aspects of the study of the effectiveness of a readability formula. The research shed light on factors that might affect reading comprehension. Finally, it proposes some recommendations on how to institute a policy in order to improve users' compliance.

The data generated were used to determine whether we might rely on a software readability scale for evaluating the difficulty of a text document in the domain of ISPs as well as factors that increase understanding of texts or enhance difficulty in understanding texts. The study contributes to our understanding of the readability in general and offering an optimal technique to measure the readability in particular.



The thesis' research questions are:

- Can readability, as gauged by software metrics, accurately reflect the comprehensibility of ISPs? If so, how?
- Does readability have an influence on understanding ISPs? If so, how can we measure it?
- Does readability have an association with security compliance intention?
- Do demographic variables have an impact on reading comprehension?

To answer these questions, a number of tasks have been addressed as follows:

T1: Reviewed related literature and related studies. To the best of this researcher's knowledge, after looking over the related works, there has not been any research conducted to examine the effectiveness of applying readability metrics as an indicator of policy comprehensibility through a mixed methods approach, and taking into consideration both human perspective as well as software perspective.

T2: After considering outcomes of existing techniques in T1, the researcher decided to use the exploratory sequential mixed methods design and started to conduct the first of seven research phases, which was choosing eight ISPs (from a larger set of thirty-five policies) for the purpose of assessing their readability by programmatic means and human comprehension tests.

T3: After the second step, a number of experts were asked for their insights on those policy ingredients that they considered key.

T4: After considering the results of the existing technique, in T3, focus group interviews were conducted to validate the views expressed by the computer experts in the previous stage.

T5: Findings from T4 were used to create comprehension tests (in the form of Cloze tests). The reason behind adopting the Cloze test is to ensure that judgment of readability is not determined solely on a mechanical basis. These Cloze tests are key to determining the comprehensibility of policy components and underpin both the pilot and main studies.

T6: After the tasks of T5 were completed, pilot studies were conducted in order to evaluate a number of Cloze tests prior to the full-scale study.

T7: Having analysed the outcomes of step T6, the main study was ready to be addressed. The aim of performing the main study was to use a quantitative, closed-ended, questionnaire instrument to determine the efficacy of applying readability metrics as an indicator of policy comprehensibility. The survey respondents' data were then analysed in order to be compared later with results of nine automated readability formulae.

T8: After analysing the results of T7, the final phase was initiated. This compared the comprehension test outcomes against readability metrics to identify the degree of similarity and difference between the results of human and software metrics.

## **1.2. Outline of the Thesis**

The dissertation is organised into seven chapters. It was felt more suitable to divide the literature review into two chapters since it introduced two domains: 1) reviews the literature and related studies on information security (Chapter Two) and 2) describes the field of readability (Chapter Three). Chapter Four explains different types of research methods and presents the methodology adopted to explain this choice of research method. Chapter Five presents the experimental design employed in this research. Chapter Six details the research results. Chapter Seven discusses the results from the previous chapter, summarises the main findings, presents the recommendations, and suggests directions for future research.

# **CHAPTER TWO: BACKGROUND AND RELATED RESEARCH**

## **2.1. Introduction**

This chapter describes the previous work and the associated research from the review of literature on information security (IS). This explains security requirements and approaches and explores features associated with IS, for example threats, vulnerabilities, countermeasures, awareness, institutional security policy and user compliance. Also included are details of information security awareness (ISA) programmes, the possible obstacles to establishing them and how to make them effective. Subsequently, there is a discussion on how to produce a successful information security policy (ISP) and how to measure its effectiveness, followed by the chapter's summary.

## **2.2. The Need for Security**

With advances in technology and its increasing use throughout organisations, institutions are becoming more reliant on online data storage and the online exchange of information in their day-to-day activities. As the number of security breaches is also growing, the risks are greater and so the need to protect and secure confidential information from intruders is crucial.

One variety of organisation that is increasingly using digital technologies is found in the health sector, which is becoming reliant on electronic health records and the use of connected health devices. However, recent reports warn that some of these devices, including pacemakers and magnetic resonance imaging (MRI) scanners, are vulnerable to wireless attacks which have the potential to kill thousands of victims (Pauli, 2016a; Knapton, 2018; RAEng, 2018). Furthermore, patients' privacy is at risk from cyber-attacks on connected health devices which could expose patients' private information, including their location history, treatments, and their current state of health (Pauli, 2016a).

Nuclear power plants are another target of hackers, which if successful, could have severe consequences. In 2017, numerous phishing emails were sent to United States electric companies but were detected and stopped before they caused any damage (Leyden, 2017). In addition, on Christmas eve in 2015 and on 17<sup>th</sup> December 2016 there were cyber-attacks on the Ukrainian national grid which led to power losses in many houses and flats (Leyden, 2015, 2016; Pauli, 2016b).

Increasingly businesses are moving into e-commerce and many more people are buying goods online. China is the world's biggest e-commerce with online retail sales of more than \$1 trillion (Tong, 2018). OnePlus, a Chinese phone company, admitted that more than 40,000 of their online customers' credit card details were stolen (Nichols, 2018; Orłowski, 2018). In addition, the personal and financial details of more than 380,000 British airways' passengers were exposed to intruders due to a compromise of their system during two weeks in August (Claburn, 2018b; Spero, 2018).

Currently, more people are moving towards using the 'internet of things' devices such as smart home hubs because of the environmental, economic, and social advantages they offer. However, many studies have expressed concern about the threats from data sharing and the observation of individuals' lives (Hill, 2017; Claburn, 2018a; RAEng, 2018).

Alongside the increase in the use of technology is a growth in security attacks. According to Leyden (2017), the number of data records compromised worldwide in 2016 was up 86 percent on the previous year, and nearly 1.4 billion data records were exposed. In the United Kingdom alone in 2015, 90 percent of large companies reported security breaches, which is an enormous increase of 81 percent compared with the previous year. In addition, almost three-quarters of small companies reported breaches which is an increase of 60 percent since 2014 (PwC, 2015).

The enormous progress and dramatic changes in the use of technology in such a short time has increased the amount of sensitive information stored on systems. Google street view (Chau, 2017) is technology that allows internet surfers to virtually visit most of the streets in the United Kingdom, prior to which, people had no concerns about the ability of strangers' viewing their properties online. Google recently declared

new features of Google maps, including an augmented reality street view mode (Gartenberg, 2018; Thomson, 2018), which assists individuals to find easier ways to walk the streets and check their individual orientation, while comparing their location with street view data as well as a live feed from the individuals' phone or other device. Moreover, Google maps has a new feature called 'For You' which offers recommendations and suggestions specifically tailored to users by referring them to stored data of their previous habits (for example, user past reviews, likes, dislikes, ratings and saved favourites). In addition, there is a group function, which enables users to share several places with friends and together they can vote on them in real time and choose their preferences. Features like these have one thing in common, in that they all require a huge amount of data. Google street view is just one example of how technology is rapidly growing, and often faster than security policies can keep up with it.

### **2.3. Information Security Background**

Information is one of the most valuable commodities in the world today. In fact, entire industries thrive on the flow, transfer, and processing of information. As with everything else of value, it attracts people who wish to exploit this information for their own personal gain. Organisations became aware of the importance of securing their valuable assets a long time ago (Greenwald, 1999). Before the invention of the computer, there were employees whose mission was to protect paper records that were usually kept in the organisation's filing cabinets. However, since the widespread use of technology by many organisations to organise, access, and share information, this method of working has changed. Organisations have transformed from being closed environments with a mainframe computer to complex environments working on distributed networks, including the internet (Sohrabi Safa, Von Solms and Furnell, 2016). This has changed the information assets of organisations into electronic forms, which are processed by information systems and communicated through the internet.

Certainly, computers, networks and the internet are fast and cost-effective methods of accomplishing work as well as making the sharing of information easier than ever (Huff and Munro, 1985). The internet has come to be a conduit for services, applications, information, content and opportunities for people and institutions.

Nevertheless, the anecdotal and empirical signs indicate that the number and seriousness of IS breaches is increasing (Sohrabi Safa, Von Solms and Furnell, 2016). Consequently, information security is required since the technology applicable to information has risks (Blakley, McDermott and Geer, 2002).

It is worth mentioning that technology alone cannot assure a secure environment for information because in addition to the technological aspects, the human aspects of IS have to be considered (Sohrabi Safa, Von Solms and Furnell, 2016). For example, in 2007, Nationwide, the largest United Kingdom financial service provider, had a security incident when a company laptop computer was stolen from a staff member's home. The laptop included private information about thousands of their customers. In this case, the Nationwide's rules had been breached by its employee who should not have held account information at home. The incident did not go unnoticed, and the Nationwide Building Society was fined almost one million pounds (Evening Standard, 2007; Wearden, 2007)

A number of scholars (Lampson, 2004; Sasse and Flechais, 2005; Schneier, 2006) warned that security is a 'people' issue, not only a technology issue as humans are the ones who are responsible for implementing IS. Regardless how great the security system is or how hard regulations or policies are to breach, there will be a frequent risk, disturbance to IS; since technology is an instrument often utilised or abused by individuals.

Nicastro (2007) highlights how reducing the vulnerabilities of information and information systems cannot be achieved by strong technology alone. Individuals, processes and technology are all essential to achieving successful IS. Consequently, institutions will be safeguarded from most threats when they are armed with the appropriate security technology, when individuals are equipped with knowledge and when there are documented processes.

## **2.4. Information Security**

Security is an enormous and broad subject as Wood (1984) acknowledges, as 'it includes the physical security of buildings, fire protection, software and hardware, personnel policies and financial audit and control' (p. 9). Security is a collection of

features and services that deal with a number of security requirements by tackling a set of incidents. As Von Solms (1999) clarifies, information security is a multi-dimensional area and an institution has to consider all aspects to protect their information assets and therefore be a secure environment.

Information security is about more than protecting an organisation's valuable resources against an intruder. Organisations need to understand that any type of attack could lead to the loss of their assets and their reputation. Therefore, successful organisations are aware of the importance of protecting their information assets to preserve a maximum level of IS (Waly, Tassabehji and Kamala, 2012). A related point to consider is that institutions are different in terms of the level of security they employ, and they vary with regards to their security needs (Garbars, 2002).

Anderson (2003) asserts that organisations should safeguard their assets, measuring their security requirements against the related risks via suitable controls while maintaining cost effectiveness. Thus, an understanding of information security is essential to enable organisations consider everything discussed. The international IS standard, ISO/IEC27001, defines information security as 'the protection of information from a wide range of threats to ensure business continuity, minimise business risk and maximise return on investments and business opportunities' (2001, p. 30).

## **2.5. Information Security Principles**

Much of the research into IS emphasises protecting and preserving the centre of IS, the CIA triad: confidentiality, integrity, and availability (CIA) (Denning, 1999; Stoneburner, Goguen and Feringa, 2002; Schneier, 2004; Pfleeger and Pfleeger, 2006; Gollmann, 2010; Waly, Tassabehji and Kamala, 2012). These three elements should be applied to a system to preserve high levels of security. *Confidentiality* ensures that the disclosure of information is restricted to authorised parties and prevents unauthorised disclosure (Thomas, 2005; Andress, 2014). *Integrity* aims to protect information from unauthorised amendment or deletion. It maintains the data, ensures it remains accurate and prevents modification by unauthorised users. Maintaining integrity is important so institutions are required to prevent illegal access as well as reverse authorised changes as required (Whitman and Mattord, 2011; Andress, 2014).

*Availability* ensures that information is available when a user requests it. It guarantees that the information remains accessible to authorised parties only in a timely and reliable manner, without affecting productivity (Whitman and Mattord, 2011; Andress, 2014).

Dhillon and Backhouse (2000a) point out that the CIA model of information security is not adequate to address IS, as it is more appropriate to information that is viewed as data held on a PC system. They propose an additional four elements, which are responsibility, integrity, trust and ethicality (RITE). These extra standards are associated with staff in an institution and are the early stages in securing an institution's information assets. The RITE term is defined as follows (Dhillon and Backhouse, 2000a):

*Responsibility* (and knowledge of roles): it is anticipated that users develop their own work practices based on a clear grasp of their obligations. *Integrity* (as an obligation of membership): as information is of the greatest value to an organisation, the organisation is required to think about how to maintain and retain integrity because the integrity of individuals could alter. *Trust* (as distinct from control): a user in an organisation has less influence on external controls and supervision and more on self-control and duty, so there should be common frameworks of trust. *Ethicality* (as opposed to rules): the ethical content of informal norms and behaviour.

The securing of an institution's information assets is a precaution against any threats whether it is an attack on or attempt to access the information. Workman et al. (2008) remark that there are three possible incidents related to IS violation, which are failure in confidentiality, integrity loss and loss of availability. These breaches are a consequence of threats and probably going to disturb the everyday functioning of an institution.

The different kinds of security threats and vulnerabilities that institutions might face are explained in the following section.

## **2.6. Security Approaches**

Some experts divide IS approaches into two main types: technical and non-technical solutions.



### **2.6.1. Technical Approaches**

Many organisations consider the adoption of technical approaches as a top priority. A number of IS studies such as (Markotten, 2002; Besnard and Arief, 2004; Vroom and Von Solms, 2004) are in favour of technical strategies. These studies highlight that using technology is an effective way of eliminating any potential security threats to systems.

Scholars in the subject of authentication mechanisms emphasise that security issues occur when an unauthorised person obtains illegal access to information system assets or when an authorised person exceeds their legitimate access level to a secure system. Thus, protecting the system from being abused should begin first by identifying all system users and then authenticating them before granting them use of the system. Authentication proves that the person is who they claim to be. Schneier (2004) supports this approach by stating, ‘no matter what kind of computer security system you use, the first step is often identification and authentication’ (p. 135). This suggests that the identification and authentication of system users are priority steps to protect any computer security system from being misused.

There are many techniques used to authenticate individuals, who attempt to access system resources. Schneier (2004) outlines these as:

- Knowledge-based approach: this is the traditional authentication approach, which is something that a user recognises/knows, for instance, a password.
- Biometrics approach: a user is an authenticator and uses, for instance, voiceprints, fingerprints, etc.
- Access token-based approach: something that an individual owns, for instance, a security token, smart card, etc.

Table 2.1 shows the advantages and disadvantages of passwords, biometrics and token authenticators (O’Gorman, 2003; Schneier, 2004)

<b>Authenticators</b>	<b>Advantages</b>	<b>Disadvantages</b>
Passwords (What you know)	<ul style="list-style-type: none"> <li>• Common and traditional approach to authentication</li> <li>• Convenient and cheaper than other authenticators</li> <li>• Has a higher key space than other authenticators</li> </ul>	<ul style="list-style-type: none"> <li>• Problem with remembering multiple passwords</li> <li>• Based on an oxymoron (has a random string which is difficult to remember but if it is easy to remember then it is non-random)</li> <li>• Vulnerable to dictionary attacks, dictionary words attack and social engineering attacks</li> </ul>
Biometrics (What you have)	<ul style="list-style-type: none"> <li>• More difficult to lend or get stolen than other authenticators</li> <li>• Strong defence against repudiation</li> </ul>	<ul style="list-style-type: none"> <li>• Serious problem with handling the stealing biometric</li> <li>• An expensive authenticator</li> <li>• May produce false positives or false negatives</li> </ul>
Tokens (Who you are)	<ul style="list-style-type: none"> <li>• Creates Multiple passcodes</li> <li>• Avoids loss of passwords</li> <li>• Better than other authenticators in preventing denial-of-service attacks</li> </ul>	<ul style="list-style-type: none"> <li>• Inconvenience</li> <li>• More expensive than passwords method</li> </ul>

Table 2. 1: The advantages and disadvantages of authenticators

Depending on the sensitivity of information, authenticator techniques can be used as either one approach or a combination of approaches to protect information. Nevertheless, it seems that the knowledge-based approach is more often used because it is cheaper than the other approaches. In addition, Schneier (2004) says that using the usernames and passwords method is the most commonly used method of the knowledge-based approaches. Thomas (2004) points out that users should be made aware of the characteristics of a strong password. For example, a) a password should not be less than eight characters, b) a password should contain a variety of classes, including uppercase letters, lowercase letters, numerals and non-alphanumeric

characters, c) a password should not contain personal information, and d) a user should avoid choosing a password made from incomplete words or slang, dialect and jargon.

Enforcing employees to create unpredictable passwords or changing their passwords periodically is an excellent way to secure organisations' information systems; however, doing such may raise the issue of not remembering passwords. This might lead users to write their passwords down to remember them, in particular, when they are infrequent users of systems, or when there is no reset password service.

A number of papers (Sasse, Brostoff and Weirich, 2001; Brown *et al.*, 2004; Carstens *et al.*, 2004; Yan *et al.*, 2004) were interested in the usability of passwords with strong password requirements and discovered that there is a lack of consideration of the limitations of the human memory for establishing strong passwords. An issue associated with this limitation is that users might write their passwords down especially if they are required to remember different passwords for each system, they are required to access.

Sasse, Brostoff and Weirich (2001) discovered that on average, an individual is required to remember 16 passwords and that the majority of login issues were caused either by people forgetting or using an incorrect password. The researchers said that the solution to this problem was to design a single sign-on login system. The main aim of this system is to reduce the load on people's memory; furthermore, it eventually encourages them to create unpredictable passwords. They also recommend that institutions increase users' awareness against security risks through education and motivation.

Other papers in the area of authentication such as that of De Angeli et al. (2005) looked at the use of passwords in non-alphanumeric forms. The researchers focused particularly on the use of graphical passwords. Their papers conclude that graphical passwords could be an effective solution to the problem of forgetting passwords, which represents a significant weakness in traditional knowledge-based authentication methods. Based on an experiment of 25 participants, they found that users are twice as likely to memorise graphical passwords than text-based passwords (Agarwal, Singh and Shukla, 2010). However, guidelines for implementing graphical passwords might be required, as a poorly chosen image can be just as readily forgotten as a poorly

chosen text-based password. Besides, according to a study by Brostoff and Sasse (2000) who were interested in testing the usability of graphical passwords, they found that graphical passwords require more time and effort to execute than alphanumerical ones.

Suo, Zhu, and Owen (2005) reveal that although graphical passwords are secure from brute-force and dictionary attacks, they are exposed to other kinds of attacks, for instance, shoulder surfing, which refers to using straightforward observation techniques, such as looking over somebody's shoulder, to obtain passwords, personal identification numbers (PIN) and other sensitive information (Kumar *et al.*, 2007). A brute-force attack is an attempt by an attacker to determine a cryptographic password by using different combinations of keys (Schneier, 2004). A dictionary attack is a method used to break a password-protected computer in which an attacker systematically tests all words in a dictionary beginning with the most commonly used passwords (Rouse, 2005).

Some researchers such as Gross and Rosson (2007) suggest that a system with strong authentication is still vulnerable. They assert that authentication alone is not enough to protect information from being accessed by unauthorised users, as social engineering still exists. Social engineering is a kind of non-technical intrusion that mainly depends on human interactions such as persuading people to do what an attacker wants, by providing them with private information or breaking the usual security measures (Schneier, 2004). There are numerous examples of social engineering and most of them are carried out over the telephone because it is usually difficult to catch the caller.

So long as security attacks exist that cannot be defeated by purely technical means, then technical approaches alone are insufficient. A lot of the research expected that the number of security breaches would reduce when technical approaches were used. But critics argue that IS programmes should also consider how the institution and its users, processes and technologies interact, and how cultural factors, human factors and architecture helps or hinders the capability of the institution to safeguard information and to manage risk (ISACA, 2009). Nonetheless, Siponen and Oinas-Kukkonen (2007) believe that some types of security threats can only be defeated by technical solutions.

### 2.6.2. Non-technical Approaches

As technology advances, there are increasing concerns about the frequency of incidents where an individual's behaviour has caused breaches in security (Besnard and Arief, 2004). Thus, users' behaviour remains the weakest link in the IS chain. This has led some researchers to study this matter with the implication being that security equipment alone does not work except that users make it work (Wagner and Brooke, 2007). Furnell (2007) indicates that peoples' attitudes and the absence of security awareness are the most significant contributors to security attack occurrences. PricewaterhouseCoopers (PwC) conducted a survey and found that human failing is the main reason behind many security incidents, rather than technology (2013). There are two ways of raising individuals' security awareness, as follows:

**A) The Punishment Approach:** is a way to influence a user to comply with regulations and rules of security as well as minimising individual-related faults in IS (Abawajy, Thatcher and Kim, 2008). Prior to applying this approach, an ISP should be established, there should be distributed campaigning and advertising, and individuals should be guided to use information systems properly before making a fair judgement of users' behaviour. As Thomas (2004) notes, having a security policy is crucial for an institution as it can protect and secure the institution's information system. Providing a policy to employees can clarify and define the acceptable and appropriate behaviour within an institution, for reducing the amount of user's misuse. However, this does not mean that all users will necessarily comply with policy (Von Solms and Von Solms, 2004). The security policy should represent an institution's IS philosophy and commitment. Moreover, the security policy should be tested and re-evaluated periodically. It should be noted that educating individuals about security issues or their misuse should take precedence over using punishment when users do not adhere to the policy. For instance, explaining the importance of complying with the security policy and recognising risks caused by user misuse, such as disclosing or writing down passwords, using pirated software, downloading unauthorised documents, accessing or modifying computerised data without permission, or exchanging sensitive information in a non-encrypted format. As stated in Wiant's paper (2005), a security policy alone may be insufficient to reduce incidents of

computer abuse because it might form the foundation on which the attack can take place.

**B) The Non-punishment Approach:** is another way to make users comply with security regulations and policies by changing users' behaviour without the threat of punishment. Some researchers divide this approach into two parts. First, by promoting security awareness through education and training (Olusegun and Ithnin, 2013). Second, by designing a usable security technology (Zurko, 2005).

**Information Security Practice Promotion:** it is crucial to promote security practice, because as Gross and Rosson (2007) emphasise, in some fields like law, finance and medicine, users are often responsible for valuable data due to their role in the institution, so if a user takes an unacceptable action, it causes a security breach that affects the entire institution. Improving employee security skills and practices is one of the most effective methods of reducing security breaches in organisations, thereby, enabling them to perform in a sensible and secure manner. In Gross and Rosson's study (2007) they discovered what knowledge individuals have about security and risks, and how users manage their security concerns. The outcomes reveal that participants recognise the value of information they deal with and perform good security practices, such as locking their screens. However, their knowledge of security-related technical components, for example, virus scanners and firewalls is limited. Furthermore, employees expect their Information Technology teams to be mainly responsible for security management. Ultimately, the researchers suggest that users should be informed about their institution's expectations regarding security and be given guidance to practice them. Organisations should familiarise all their employees with general security principles and ensure that those employees who deal with sensitive data are experts in security because they might access the organisation's network remotely from their own computers, which may be vulnerable to attacks.

Based on the experiments of Bakhshi, Papadaki and Furnell (2009), individual awareness and experience has an effective impact on a user's ability to protect themselves from different security threats. People who have had experience of social engineering are more careful about sharing personal information than those who have not. Organisations should encourage their staff to become up-to-date with security

knowledge. To this end, an organisation can provide monetary incentives to its employees to attend security-related workshops or courses. A 2004 survey regarding IS breaches shows that large organisations (those with more than 500 employees) suffer more from staff-related breaches than from attacks from outsiders. In addition, the report points out that most of the staff-related incidents were because users neglected or violated security policy (Vroom and Von Solms, 2004).

**Usable Security:** scholars in this field stress that the human-computer interface performs a significant part in security issues. Maxion and Reeder (2005) caution that security might be compromised when individuals make errors due to the user interface. Thus, improved interface design may prevent human error. Markkotten (2002) believes that security tools are worthless if user interfaces are difficult for people to use. He suggests that management of the user interface has to remain with the individual; however, the system has to be designed to be as easy to use as possible to guarantee the required security level. In contrast, others note that it is better for users to be aware of the consequences of their actions' prior to letting them make decisions about security issues. For instance, a number of users do not make any effort to look at the digital certificate that is provided by a browser and if they do, the individual does not usually have enough knowledge to judge whether to trust the certificate or not (Zurko *et al.*, 2002).

## **2.7. Rewards or Sanctions Approach**

There is debate about which approach should be adopted, punishment or non-punishment. A number of organisations use rewards and sanctions as motivation and deterrence. Some researchers support the punishment approach, which is called the negative enforcement strategy; however, a number of scholars advocate the non-punishment approach (reward) which is called the positive enforcement strategy. The punishment approach has a number of positive effects, such as deterring deviant acts and decreasing computer abuse incidents (Trevino, 1992). Not only should the punishment fit the misbehaviour, but also the punishments should be pre-defined. This can be indicated in the organisation's regulations and policies and distributed to each employee. The organisation should use the following sequence: the creation of a

security policy, campaigning and advertising, training, punishment and reward, and evaluating and readjustment (Marks and Rezgui, 2009).

Some researchers claim that the punishment strategy has more negative consequences than benefits, for instance, anxiety, aggressive acts, or withdrawal. Nevertheless, other empirical studies suggest that these unwanted side effects are especially weak and may only occur if the punishment is imposed indiscriminately (Chen, Ramamurthy and Wen, 2012). Users should be aware of the punishment sanctions and the certainty of being punished when they violate the security policy (Straub, 1990). An example of deterrence could be in the form of a reduced wage, increased monitoring, stricter control or criticism. A number of studies are in favour of the sanction approach only as they claim that the reward approach has no significant effect on users' adherence to the ISP (Puhakainen, 2006; Siponen, Pahnla and Mahmood, 2010; Siponen, Adam Mahmood and Pahnla, 2014).

On the non-punishment side, there are various examples of possible rewards such as bonuses, certificates, wage increases, praise and recognition in the form of titles, parents' insurance coverage, extended breaks, or a show of appreciation. There are many benefits of adopting the reward approach which include: improving employees' performance, raising job satisfaction, directing users' behaviour, motivating and retaining talent, and encouraging excellence (Chen, Ramamurthy and Wen, 2012). In a study investigating the effectiveness of the reward and sanction procedure, Gonzalez and Sawicka (2002) report that users believe reward to be the greatest encouragement to comply with institutional policies; yet, if or once the reward is removed, users' might not adhere to the policy. An empirical investigation conducted by Siponen et al. (2010) shows that the non-punishment approach has no impact on users' compliance with ISP, whereas, the punishment approach has a significant effect on users' adherence to the ISP. A study by Boss and Kirsch (2007) reveals that rewards do not have a significant impact on users' compliance with ISP.

Furthermore, some critics have doubts about the effectiveness of the reward strategy for the following reasons (Kohn, 1993; Amabile, 1998). First, rewards might only assist the progress of temporary compliance. Secondly, positive stimuli may produce unexpected consequences. For instance, generating a controlling work



environment, especially when staff members feel managers are controlling them. Thirdly, this strategy can go wrong if the organisation considers other factors such as the number of times a user has been awarded. More specifically, although a person deserves to be rewarded, the organisation may refuse to nominate him because he has already previously received a reward. Finally, when the rewards are given based on personal relationships not the employee's performance.

In conclusion, even though the choice between punishment and non-punishment approaches has long been an important issue for organisations, researchers do not agree on the relative merits of each approach. Institutions need a balance between productivity and strict security. Therefore, we can say that motivational factors with respect to the act of adherence to ISP are as significant as sanctions or preventive elements in efficient IS management.

The question is whether adopting a combination of the use of punishment and reward strategies can bring the benefits of both approaches regardless of the drawbacks of each. Many studies found that adopting both procedures is fair to employer and employee, and could have a significant impact on compliance with the ISP (Bulgurcu, Cavusoglu and Benbasat, 2009, 2010; Chen, Ramamurthy and Wen, 2012).

## **2.8. Information Security: Threats, Vulnerabilities and Countermeasures**

Every new decade brings new threats to organisations. Security threats are any circumstances that create the possibility of loss of or harm to IS (Amoroso, 1994; Pfleeger and Pfleeger, 2006). There are so many types of threats to businesses that the list seems never ending. However, a number of publications and surveys have enumerated the diversity of threats to IS that an organisation might face (for example (Fulford and Doherty, 2003; Schou and Trimmer, 2004; Al-Awadi and Renaud, 2007; Whitman and Mattord, 2011; PwC, 2015; Ernst and Young, 2018)). The following threats, which are classified as external and internal, are identified in the studies previously referred to:

- **External threats:** (a) viruses, worms, spyware and Trojan horses, (b) natural disaster, (c) spam, (d) hacking attacks.

- **Internal threats:** (a) installation /use of unauthorised hardware, peripherals or software, (b) misuse of computer access controls, (d) physical stealing of hardware /software, (e) human mistake as violation, (f) deliberate damage by displeased users, (g) using an organisation's facilities for illegal communications or activities.

A number of researchers (Sterling, 2012; Marcotte, 2013; Krebs, 2014; Georgetown University, 2015; Mano Ten Napel and Novealthy, 2015) highlight how attackers always find new methods to forge, steal or gain access to systems. They mention that the top threats to IS in this century are as follows:

- **Technology with weak security:** many new tools and devices, such as most Internet of Things devices, are able to access the internet but they are not secure by design. Every unsecured connection represents vulnerability. The fast growth in technology is a testament to innovators, but there is a serious delay in terms of security.
- **Social media attacks:** attackers may get leverage from social media as a way to hunt for victims. Cybercriminals use many ways to target victims such as creating fake accounts, celebrity name misuse, and spreading spam or malware.
- **Mobile malware:** security specialists have raised security concerns since the early stages of mobile devices and their connectivity to the internet. A study by the University of Cambridge revealed that nearly 90 percent of Android smartphones are exposed to at least one critical vulnerability (Thomas, Beresford and Rice, 2015). Apple products are not immune either, and approximately 40 iPhone operating system apps were pulled from Apple's official app store in 2015 because they were infected with malicious code (Kochetkova, 2015; Lovejoy, 2015).
- **Third-party entry:** also known as a supply chain or value chain attack, which happens when someone infiltrates your system over an outside partner or provider and gains access to a range of valuable and sensitive information. Organisations require concentration on the weakest spots in their supply chains. A recent survey found that 56 percent of organisations have

experienced a data breach caused by one of their vendors (Ponemon Institute, 2017).

- **Outdated security software:** there is a serious risk posed by outdated software, browsers, and operating systems because it will be vulnerable to any new malicious code.
- **Social engineering:** this is not a method run by cybercriminals, but it is an unpredictable and effective form of intrusion as it depends on human interactions and psychological manipulation to get the desired information.
- **Lack of encryption:** encryption is important to save data from breaches. Many countries are taking firm action on this issue by imposing very high fines on any organisations that fail to meet security standards, including data encryption. For example, two companies in the United States of America were fined almost two million dollars in 2014 due to lack of compliance with the Health Insurance Portability and Accountability Act (HIPAA), including stolen, and unencrypted laptops (Gold, 2014; OCR, 2014).
- **Corporate data on personal devices:** storing or transferring corporate data should be limited to business-only devices and confidential data should not be accessed on personal devices. Ignoring that could pose a data security risk to the people or the organisation.

These threats can also cause security breaches, for instance, spam email is a type of threat, which becomes a breach of security when a user opens the spam email.

## **2.9. Information Security and the Institution**

According to Schneier (2001), institutions should know that it is not possible to have a completely impenetrable security system regardless of their resources. Therefore, institutions should look for realistic security to avoid any serious breaches which could cause the institution to collapse due to litigation, harm to brands, financial loss, loss of client confidence, etc.(Rainer Jr *et al.*, 2007).

Currently, many institutions are heavily reliant on information technology to share information and other resources for completing their work (Dhillon and Backhouse, 1996). They face growing risks because they upload their valued assets into an online

database, which makes it accessible from a local network or via the internet and could lead to intruders gaining access to it (Dourish *et al.*, 2004).

There are a number of hindrances to implementing IS in an institution and the first one is complacency. Dhillon (2001) believes that employees' complacency is one of the main reasons for accidents and breaches of security in an institution. Gritzalis (1997) states that complacency arises in users when they become overconfident regarding the institutions' security. Hence, when the institutions' security systems are running efficiently without disturbance, complacency might occur. Hentea *et al.* (2006) assert that awareness and education are effective ways to reduce user complacency.

The second IS problem is security planning. Fitzgerald (2016) states that regardless of the size of the institution, security planning is very important to improve the security of the entity. O'Connor (1993) names three interrelated elements with regard to security planning: strategic, tactical and operational plans. Tryfonas *et al.* (2001) explains strategic planning as what the institution would like to use, tactical planning as the methods and techniques to be employed throughout the operational planning, and operational planning as the development of specific comprehensive plans for any project.

There is one other issue related to ISP, which will be covered in detail later, but it is worth mentioning that, there is no 'magic bullet' for achieving perfect security, even with money or time (Siegel, Sagalow and Serritella, 2002). In terms of the money factor, it plays an imperative role in the choice of security measures and a number of institutions are turning to technology to assist them in shoring up their protection, nonetheless, IS issues cannot be managed solely from a technical aspect (Posthumus and Von Solms, 2004). Taipale (2004) stresses that no one can ignore how important the technical perspective is but that it cannot provide absolute security.

## **2.10. Information Security in Higher Education**

Higher education institutions face more challenges than some other organisations because of 1) the complexity of their information systems, 2) the thousands of new students that join the universities every year, and 3) because they provide open access

to their constituents and to the public (Cox, Connolly and Currall, 2001; Rezmierski, Seese Jr and Clair II, 2002; Katz, 2005; Burd *et al.*, 2006; Rezgui and Marks, 2008). Rezmierski *et al.* (2002) point out that ‘the nature of the teaching-learning environment requires a vast array of operating systems, platforms and networks to meet the needs of various scholars thus increasing the complexity of the overall system’ (p. 577). Rezgui and Marks (2008) state that colleges and universities are targets for attackers as a result of the high computing power they have and the continuous open access granted to a wide range of user types both inside and outside the institution. This type of sector usually contains students, faculties, staff, administrators, labourers, and so forth. There are many different campuses with various network resources, where employees and students, for instance, expect to access data or their own documents from classrooms, laboratories, libraries and off campus. Universities’ faculties generally comprise different departments where the security requirements differ from department to department and from faculty to faculty.

Katz (2005) notes that the higher education sector is exposed to various possible human threats, especially to IS, because of the quantity of confidential information being dealt with. Rezgui and Marks (2008) emphasise that universities and colleges are amongst the least secured information entities and a small number of them provide security awareness training, despite being prone to many threats and attacks. Previous studies support what Rezgui and Marks (2008) found in that one-third of surveyed organisations provided security awareness training for employees and students (Caruso, 2003; North, George and North, 2006).

## **2.11. Raising Awareness**

With advances in technology and its increasing use across organisations, the demand to protect confidential information has become the requirement of the age. This need has given birth to ISA programmes in corporations worldwide. Many studies indicate that the attitudes of people and their lack of security awareness issues are the most notable contributors to security incidents (Furnell, 2007). Furthermore, previous studies also reveal that the weakest links in the information security chain are individuals, which is why ISA is considered very important (Lebek *et al.*, 2013). To reduce the incidence and severity of such attacks, it is necessary to raise the level of

ISA within organisations and even in the public (ALArifi, Tootell and Hyland, 2012). In fact, there are a number of websites, blogs and e-learning courses, under the supervision of reputable institutions in IS, that claim to offer resources to develop an engaging and high-impact security awareness programme (SANS, 2010; QA Ltd, 2014; Leicester University, 2016; Wombat security, 2017; *InfoSec skills*, 2018).

Such programmes and platforms are created to assist organisations to build a culture of security within their entities. The e-learning courses are developed to help users develop their perspective on IS. Johnson (2006) states that the target audiences vary depending on the security awareness needs of an institution. It might include all or some of the following: employees, line management, executive management and boards of directors, field employees, laptop users, IT sector, email users, sellers and suppliers.

Institutions should use both educational/interactive methods and informational methods with some creativity and diversity in the choice of means to help individuals recognise the potential threats and understand how to respond. Educational/interactive methods include using slide presentations, training, brief targeted sessions, online learning modules, demonstrations, videos, workshops, and so forth. Informational methods include adopting leaflets, short articles or news stories, intranet security website postings, e-mail warnings, information security guides, tips-of-the-month, flash cards, newsletters, etc. Institutions might also employ promotional methods such as, events/fairs, screensavers, intranet banners, page to the security page, posters, puzzles and games, pre-printed notepads or sticky notes, T-shirts, mugs and cups, mouse pads, and stickers for raising awareness of security threats among workforce populations (Cox, Connolly and Currall, 2001; Johnson, 2006; Khan *et al.*, 2011; Awawdeh and Tubaishat, 2014).

Despite the methods suggested to raise security awareness, their usefulness cannot be evaluated. It is questionable whether users notice certain messages posted in different places within an organisation. However, there are processes to assist in identifying exposure to messages, if not the effectiveness of the messages themselves. Institutions might compare the rate of attacks before and after deploying such methods to evaluate their effectiveness as well as helping to increase security awareness.

Nevertheless, this is reliant on several complex factors and does not necessarily provide much insight into the effectiveness of security awareness methods. Kajzer et al. (2014) used a survey to assess the effectiveness of security awareness messages and found that the usefulness of such methods varies depending on user personality, and is not always as one would anticipate. They argue that some security methods appear helpful to security efforts based on certain personality traits, while other personality traits make the user less interested in certain methods and therefore, security messages can be counterproductive in achieving the intended effect.

Craig (2011) believes that security awareness systems may collapse if they are seen to be quick, simple, inexpensive or knee-jerk reactions to a current threat. He makes a number of points that organisations should be aware of: A) lack of widespread support for the reasons for the security awareness (SA) programme if run by one group (e.g. the information technology department), B) the SA programme lacks credibility, relevant skills, development and recent information because the wrong person is in charge, C) a lack of budget and resources to support the programme, D) a failure to deliver the SA programme's objectives in an effective manner, E) no evaluation of the SA programme regarding whether it has been effective or not in both the short and long-term, and F) the SA programme does not address the business needs and is driven by availability, content or management directive.

According to Craig (2011), there are a number of pitfalls that can undermine the SA programme even with the right timing and planning, these include: 1) disagreement about where the funding should come from, 2) an unacceptable one size fits all approach, 3) target users are too limited, 4) failure to set a strategy for the purpose of maintaining and developing the programme, 5) poor feedback mechanisms about problems, successes and queries, 6) limited or ineffective support from the senior management, 7) no sanction policies or procedures to maintain the SA programme, 8) lack of integration of programme components, 9) lack of security behaviour of many employees (e.g. carelessness, user security errors), and 10) poor novelty value.

## **2.12.The Importance of Information Security Awareness Programmes**

Information is one of the most valuable commodities in the world today. In fact, entire industries thrive on the flow, transfer and processing of information. As with everything else of value, this attracts people who wish to steal and exploit this information for their own personal gain. Khan et al. (2011) explain ISA as guaranteeing that users are aware of the institution's rules and regulations. Threats to information usually arise in the form of malware attacks, hacking attempts, denial of service (DoS) attacks, etc. and such attacks often involve an unsuspecting human agent whose actions introduce the attackers to the system (ALArifi, Tootell and Hyland, 2012).

The primary aim of ISA is to ensure that there is no loss of business or any type of liability for the organisation due to the loss of information. Additionally, organisations have an ethical as well as a legal responsibility to ensure that their confidential information is protected from malicious access (Chan and Mubarak, 2012).

Studies indicate that employees are the weakest links in the information security chain, which is why ISA is considered important (Lebek *et al.*, 2013). One way to reduce the incidence and severity of incidents is to raise the level of ISA within organisations and the public (ALArifi, Tootell and Hyland, 2012). Although ISPs and procedures are routine in most organisations, many people ignore such precautions. One potential cause of such behaviour may be reduced if institutional ISPs are specific and clear to those who are required to comply with them. Furnell and Clarke (2012) state that information security behaviour can only be accepted once it is combined with technological aspects.

To ensure minimum losses and the safety of online data, institutions are working to make ISA their primary concern, particularly in the financial and banking sector, where organisations present their employees with guidelines for the protection of corporate data. General security awareness, organisational budgets and the level of employee computer skills are noted as major barriers to increasing ISA (Shaw *et al.*, 2009). The main benefit of ISA is to ensure that there is no loss of business or any type of liability for the organisation due to the loss of information. Additionally,



organisations have an ethical as well as a legal responsibility to ensure that their confidential information is protected from malicious users (Chan and Mubarak, 2012).

### **2.13.Obstacles to Establishing Information Security Awareness Programme**

There are a number of obstacles that organisations might face when trying to implement ISA programmes. These include lack of: funding for the ISA programme, training, defined roles and responsibilities, policy implementation, communication and documentation, employee involvement, risk management, employee awareness, reward and sanction, employees' security perceptions, goals, vision and structure. There are many influences affecting ISA, for instance: language, management commitment, organisational culture, business policies and procedures, security policy, security standards, information resources, risk analysis and risk assessment, organisational culture, IS users' morality and ethics, motivation, users' computer skills, social norms, social norms and human behaviour.

There are also behavioural obstacles, for example; employee habits, it being someone else's problem, a lack of monitoring and consequences, lack of management or colleague support, individual beliefs and attitudes. Furthermore, there are obstacles such as; security awareness training, complexity, information overload, lack of reinforcement, not enough differential in learning styles, insufficient knowledge sharing and discussion, too much technical jargon, and not enough convincing and persuading. There are of course, other factors that enhance the success of the ISA programme such as involvement, motivation, awareness, communication, reinforcement, satisfaction, convening and persuading, assessment and positive feedback (Abawajy, Thatcher and Kim, 2008; Tsohou *et al.*, 2008; Boujettif and Wang, 2010; ENISA, 2010; Alzamil, 2012; Rastogi, 2012).

Nevertheless, organisations can overcome some of the most common obstacles when they are aware of them before or during the planning and implementation of the ISA programme.

## **2.14. The Effectiveness of Information Security Awareness Programme**

When the ISA programme is established, it should be assessed periodically, and its effectiveness evaluated. There are a number of methods for doing that, such as:

- **Feedback:** users should be encouraged to express their feelings about the programme and about what they do or do not prefer about the security awareness training, awareness activities, and what they might prefer to be done instead (e.g. feedback channel).
- **Metrics:** it is important to choose the right metrics for an institution to obtain a better result and convince the budget holder of the importance of the SA programme.
- **Value it with others:** the SA programme could be assessed by other teams within the institution. For example, by requesting the human resources (HR) officers to provide reports about security breaches before and after the SA programme is established (Wilson and Hash, 2003; Craig, 2011).

There are many advantages of establishing ISA programmes in an organisation, such as a) they mitigate the overall security risk, b) they improve the confidence of users, managers and board members in the organisation's security, c) improving reliability and accuracy of an organisation's information, d) reducing internal incidents, mistakes and omissions, e) enhancing the discovery of remaining security events, f) the confidentiality of critical information is better protected, g) it saves the organisation's money by reducing the number of security breaches, h) it raises employee morale and improves productivity, i) it increases compliance with regulations and laws, and j) compliance with the CIA triad, privacy and security standards. In addition, establishing ISA programmes leads to the involvement of many functions and departments in an organisation, for instance, executive management, help desk, safety department, and public relations and thus creates a culture of security awareness (Johnson, 2006; ENISA, 2010).

However, organisations should be aware that the ISA programme is expensive at the outset, as it includes the cost of salaries for the security awareness team or

coordinator and money for materials, training and room rentals. In addition, the ISA programme has indirect costs such as the cost of time spent by the intended audience on training and the time spent by sectors associated with promoting the ISA programme (Johnson, 2006). Each institution has to discover the right balance for them as there is no 'one size fits all' solution, but making the approach simple is likely to maintain its cost-effectiveness (ENISA, 2007).

## **2.15. Programme Drivers**

Institutions usually use three popular programmes to address security awareness; requirements-driven programme, means-driven programme and needs-driven programme. Many organisations prefer to use a combination of two or all three programmes, but the majority use one (Roper, Fischer and Grau, 2005).

In a requirements-driven programme, an outside authority imposes a number of requirements such as standards and regulations that must be fulfilled by the programme developer. In the government sector, security education requirements are included as security regulations, whereas in the industry sector, an organisation has to adhere to the law or regulations' requirements that cover the organisation's scope.

The means-driven programme depends on the available resources in an organisation such as materials or people. In other words, the institution itself addresses what they have and attempts to deal with it. Many existing security education programmes adopt this kind of programme due to insufficient money or time. Roper, Grau and Fischer (2005) describe an example of a programme driven by the availability of a particular person and certain materials in one large institution with many different assets to protect. The individual in question was a good public speaker, an excellent writer, exhibited clear and prominent security posters about security programmes, and the reasons for them. This institution's security education programme mostly dealt with personnel security and surety, which is just one aspect of a full security programme and was due to the person running it being responsible for personnel security, which despite being well written and presented only covered a small part of the institution's security requirements.

In a needs-driven programme, the security members examine the security programme and suggest the requirements for running it. Then, they choose the most suitable employees in the institution to facilitate it and remove any obstacles to make it work. Roper, Grau and Fischer (2005) emphasise that on first impressions, many people think the needs-driven programme is ideal and should be adopted instead of a mixture of two or three programmes. However, lack of confidence is the major flaw in this programme, as the security staff members tend not to have confidence in their ability to deliver a successful security education programme.

## **2.16. What is an Information Security Policy?**

Many studies indicate that employee attitudes and lack of security awareness are the most notable contributors to security incidents (Furnell, 2007). Institutions should have security policies that include pertinent documentation that reflects local information security philosophy and commitments (Johnson, 2006).

An ISP is defined as one of the most significant components required in an institution to effectively ensure information security resources. It describes the role it plays in achieving and assisting the institution's vision and mission (Höne and Eloff, 2002b). The ISP is a set of rules or requirements, related to IS, enacted by an institution that has to be adhered to by all the users within the organisation to protect the confidentiality, integrity and availability of information and other valuable resources as well as prevent any possible security threats (Tryfonas, Kiountouzis and Poulymenakou, 2001; Canavan, 2003). According to ISO/IEC (2018) one of the objectives of the ISP is to deliver management direction and support for IS. A number of researchers note that it is those users who neglect to adhere to the ISPs within their institution who cause most of the security problems (Herath and Rao, 2009; Myrsky *et al.*, 2009). However, establishing an ISP alone is not sufficient unless it is followed by standards, procedures, guidelines and policy enforcement. In addition, institutions are required to validate their ISP and ensure it is structured and planned effectively (Von Solms and Von Solms, 2004; Vroom and Von Solms, 2004; Hong *et al.*, 2006; Alnathier and Nelson, 2009; Ammann and Sowa, 2013; Sommestad *et al.*, 2014).

As Higgins (1999) emphasises 'without a policy, security practices will be developed without clear demarcation of objectives and responsibility' (p. 1) and David

(2002) asserts, ‘security is not what you do... security is how well you adhere to your formal security policies’ (p. 506). The ambition of implementing an effective ISP will not be achieved unless users are familiar with its content and comply with its requirements. A report released by PwC reveals that there were staff-related breaches in 72 percent of institutions where security policy was poorly understood (PwC, 2015). Biddinika et al. (2016) believe that ‘readability is the starting point for the comprehension of a text’ (p. 1351). Notably, Doherty and Fulford (2003) declare a set of factors that cause users to ignore security policies, one of which is the complexity of the policy document.

## **2.17. The Role and Importance of the Information Security Policy**

ISPs often form part of an organisation’s official regulatory framework. The role of the ISP is to ensure that any decisions and actions are consistent with the objectives of an organisation. Most organisations now impose ISPs or ‘conditions of use’ agreements upon their employees. The need to ensure that employees are informed and aware of their obligations regarding information security is apparent. Less obvious is the correlation between the provision of such policies and their compliance. Although ISPs and procedures are routine in most organisations, many people ignore them. One potential cause of such behaviour may be reduced if institutional ISPs are specific and clear to those who are required to comply with them.

Policies should be considered as rules or principles that users understand and follow. To this end, ISPs have to be expressed in a manner that is received as commonplace and accepted as part of regular tasks (Ammann and Sowa, 2013). Users are frequently identified as the key vulnerability in an organisation’s information security profile and are often the main cause of security incidents. Höne and Eloff (2002b) believe that users ignore ISPs because they do not fully understand them. Hence, if users do not fully understand the content of ISPs, a major factor in security incidents may be the security policies themselves. Accordingly, authors responsible for writing an ISP should try to ensure that the information in the policy reaches its audience easily and effectively. In addition, the ISP should be written in a way that attracts the user’s attention and should not be too long or too technical. Individuals

may see the policy as a nuisance if they feel there is no association between it and their everyday duties (Höne and Eloff, 2002b).

## **2.18. Compliance with the Information Security Policy**

A crucial requirement for institutions is to have absolute compliance with the ISP to reduce violations of it. Even though ISPs and procedures are routine in most organisations, many people tend to ignore such precautions. Many organisations, especially in the financial and banking sectors, give employees clear guidelines relating to the protection of corporate data.

The user's compliance with the ISP is an important factor that should be addressed before and after the ISP is established. Scholars admit that there is insufficient empirical research regarding user's compliance with ISPs, even though it is considered a key issue for institutions (Puhakainen, 2006; Ernst and Young, 2012). There is a direct connection between the awareness level and the compliance level among users within an institution. The results from Lane's (2007) study confirm the previous statement that users' compliance with information security is enhanced once an organisation has an adequate ISA programme that is taken seriously.

It is estimated that more than half of security breaches, either intended or unintended, are created by users' not complying with security policies. The reasons for the violation of security policies are usually due to users' negligence or ignorance of them (Vance, 2010). A number of researchers such as; (Wood, 1982; Straub, 1990; Parker, 1997; Bequai, 1998; Tudor, 2000; David, 2002; Kankanhalli *et al.*, 2003), are in favour of using sanctions, grounded in deterrence theory to deal with the issue of employees' negligence of ISPs. Siponen et al. (2014) assert that staff comply more with security policies where there are sanctions. Furthermore, Puhakainen (2006) in his study of employees' failure to comply with policies, recommends to use scenario-based exercises as part of training session as a good technique to make employees realise what harm could happen if the SPs are violated . This information might be useful to those organisations who are keen to improve their employees' behaviour and successfully gain higher levels of compliance with their ISPs (Beautement and Sasse, 2009).

The attitude of users can be affected by the advantages of compliance, the disadvantages of compliance and the cost of failure to comply. The advantages of compliance are the significant benefits from the safety of resources, whilst the disadvantages are an increasing workload and work impediments. Failure to comply causes significant costs in terms of the vulnerability of resources and sanctions (Bulgurcu, Cavusoglu and Benbasat, 2010; Vance, 2010). The results of a survey carried out by Bulgurcu et al. (2009) with 464 participants reveal that ISA and perceived fairness positively influence attitudes, and consequently attitude positively influences the aim to comply.

There are a number of variables that influence compliance with ISPs and researchers have identified more than 60 factors that could affect this (Sommetstad *et al.*, 2014). A policy is worthless if an organisation believes that just by having it; employees will adhere to it, without any need to monitor it. Myyry et al. argue that users may tend to state their adherence to ISPs in general, but at the same time may disregard or violate particular rules or policies (2009). Hence, organisations should regularly review, monitor and enforce compliance with such a policy (Von Solms and Von Solms, 2004) and should act wisely if any events of non-compliance have occurred.

### **2.18.1 Threats to Compliance**

There are a number of possible reasons why users do not adhere to the ISP, as follows:

- User attitudes or personality: according to Thomson and Solms (2007), individual values and beliefs have an impact on user behaviour towards IS. Leach (2003) highlights that users might not comply with the institution's values, standards and beliefs when they do not match their own.
- Lack of knowledge and skills: users will not be capable of complying with the organisation's regulations and rules when they have deficiencies in this aspect. Moreover, users might not adhere to the ISP if they do not have enough motivation to comply (Neal and Griffin, 2002).
- Leadership: management plays a vital part in inspiring users to complete security assignments (Neal and Griffin, 2002). Von Solms and Von Solms

(2004) assert that users' negligence of an organisation's security policy and regulations could be due to management failure to train users about what they are required to do regarding IS.

- Institutional culture: this can have a positive or negative impact on the user (D'Arcy and Greene, 2014). Chang and Lin (2007) believe that the security culture refers to the values and beliefs regarding IS that are shared by users at all levels of the institution. For example, an institution's users will understand that the security issues associated with using an institution's computer are different to those when using their personal computer.
- Invisible security policy: the ease of access to and readability of the organisation's policy can affect users' security behaviour (Leach, 2003). Doherty and Fulford (2005) noticed that several institutions have deficiencies in their ISPs. In addition, Barman (2001) warns that a policy can send incorrect messages and not been taken seriously, when it is not clear enough to users.
- Work pressure: one of the aims of a security policy is to protect the information system assets while maintaining its effectiveness. The system should not be excessively secure to prevent legitimate users from obtaining the information they need to complete their work. Employees may ignore compliance with IS if they think it disrupts their work, as some employees are more concerned about completing their work (Wood, 1984). A study conducted by Awawdeh and Tubaishat (2014), reveals that a number of participants violated security policy when it delayed their work at times of pressure.

There are significant challenges to reducing security attacks. Organisations and users have to cooperate to obtain benefits for each party. Furthermore, institutions are required to think carefully and be thorough when implementing ISPs, because as Dhillon and Backhouse (2000b) emphasise, when 'facing pressures of organisational cost containment and external competition, many companies are rushing headlong into adopting information technology without carefully planning and understanding the security concerns' (p. 127).



## **2.19. How to Make Successful Information Security Policies**

There is no clear single approach to ISP design or content selection that guarantees that an organisation can accomplish its IS aims. However, the clarification of the requirements and concepts of the ISP to users is key (Höne and Eloff, 2002b). Therefore, evaluating the comprehensibility of the policy may assist them in determining whether the ISP is likely to be effective. For example, if auditors certify that controls and security measures are working in accordance with the policy, this indicates a good fit between the policy and those charged with its application. Yet, insufficient controls and security measures may produce an ineffective policy (Ammann and Sowa, 2013).

Authors should consider the writing style and the way in which the ISP is presented to users. Höne and Eloff (2002b) suggest that the ISP document should be presented in an eye-catching style to attract the users' attention and ensure the desired objectives are delivered. Notably, organisations should not leave the documentation of ISPs to technical staff in isolation from others. Although they may have experience in information security technologies, this may not be matched with experience of users' understanding and how IS may suit the broader organisational culture (op. cit.).

An ISP cannot be successful unless all users are familiar with it. Consequently, institutions should strive to distribute the ISP efficiently and be open to addressing any issues related to the transparency of its content. Additionally, the ISP should be regularly enhanced and updated to ensure it continues to fit with the institution's vision and mission. Another aspect of the required transparency is that ISP authors should consider the readability of their text, as this is fundamental to its comprehensibility and thereby its effective operation.

There is a direct connection between the effectiveness of an ISP and the effectiveness of IS. As Höne and Eloff (2002b) express 'at the end of the day, an effective ISP will directly result in effective information security' (p. 15). As Canavan (2003) indicates, setting ISP into practice is an approach of ISP enforcement. Therefore, whenever an institution puts an ISP into practice, users can follow the requirements and understand their rights and duties within an organisation. Adopting an effective security policy will help individuals to comprehend what is acceptable and

reasonable behaviour, and achieve a better environment (Höne and Eloff, 2002b; Von Solms and Von Solms, 2004). Mader and Srinivasan (2005) state that for the purpose of increasing the effectiveness of ISPs, punishments should be imposed on those who do not adhere to the policies.

Gaunt (1998) judges the success of an institutions' security policy by the clarity of the text outlining the employee's rights and responsibilities. There are numerous standards and guidelines that have highlighted the importance of the security policy (Höne and Eloff, 2002a; Fulford and Doherty, 2003; Siponen and Willison, 2009), but despite this, several scholars have discovered that many organisations in developed countries (such as, Canada, United States, and United Kingdom) have introduced ISPs but do not appear to understand their importance (Andersen, 2001; Hinde, 2002; Gupta and Hammond, 2005; Hong *et al.*, 2006). To be effective, the ISP should be a living document that is regularly enhanced, developed and updated (Briney, 2000). It is also important that the ISP is relevant to an organisation's security objectives. The ISPs must be implementable and enforceable and, in particular, distributed and communicated throughout the organisation and to each individual responsible for the security of, or with access to the institution's information assets (Höne and Eloff, 2002a). In addition, Palmer et al. (2001) suggest a number of points that should be considered to produce a better policy. These are: A) the policy's scope should characterise what assets are influenced and to whom the organisation's policy applies (e.g., all users, part-time users, consultants, or clients), B) it is necessary to have a well-structured policy with easy references to specific sections, C) the policy file should be painstakingly worded and all terms should be accurately and precisely characterised, and utilised precisely as planned, D) the policy must be technically and organisationally feasible, E) the policy should define duty, accountability, and lines of authority throughout the institution, and F) the first edition of a policy document, and subsequent updates, should contain a version number and a date.

## **2.20. Evaluating the Success and Effectiveness of Information Security Policies**

Some factors may minimise the efficacy of the ISP even before it is introduced (proactive/prior factors) while other factors may minimise the efficacy after use

(reactive/post factors). A list of relevant factors would include: readability of ISP documents, level of user awareness, ethical conduct policies, organisational culture, adoption of recognised standards, proportion of detecting viruses and unauthorised software, audit results, outcomes of users surveys, levels of user compliance, reducing lost productivity, reducing security incidents, level of user training, consistency in enforcement of ISPs and standards, senior management commitment to IS initiatives, appropriate employee education and awareness regarding information asset protection, achievement of ISP targets within the available budget, a balance of effort between achieving short-term goals and long-term targets, the extent of alignment of the ISPs with the organisation's objectives and the cost justifications for IS (Chapple, 2005; O'Bryan, Caraway and CISA, 2006; ENISA, 2007; Alnatheer, 2015).

The ease of reading<sup>1</sup> ISP documents is a 'proactive factor'. Investigating the readability of ISPs may help in increasing the compliance with regulations and rules of security, which result in the increased effectiveness of ISPs. In principle, this may be achieved by testing documents using software readability metrics and/or testing using human readers (comprehension tests). Readability metrics may help in assessing the clarity of ISPs and could offer an easy means for an organisation to gauge their own policy through self-assessment (Chapin and Akridge, 2005). Unlike other measures, readability metrics can assist in improving the inherent properties of ISPs, i.e., their textual content and comprehensibility. The supposition is that whenever the policy is not fully understood or the text's content is hard to read, then the policy will not readily be used or will be used insufficiently (Ammann and Sowa, 2013). Jayaratne (2014) makes it clear that if text is written at a higher level of comprehension than the target audience is capable of appreciating then it will probably not be understood or followed, so its purpose will not be met. Ideally, institutions need to ensure that their security policies can be understood by all employees regardless of their level of education.

Whenever a document is intended to be presented to any group of people, readability or reading ease metrics may be employed to estimate the level of the

---

<sup>1</sup> The term 'ease of reading' is used synonymously with readability

document and gauge how straightforward it is to comprehend the text. A readability index is evaluated by using statistical text analysis.

## **2.21. Summary**

In view of what has been discussed in this chapter, it is necessary for users and institutions to comprehend what makes information secure in organisations including the higher education sector. The institutions need to know what threats they face and what issues they should be aware of as well as the possible organisational vulnerabilities. Where there are existing policies, they need to know why users neglect or violate them and how a successful policy improves responsible behaviour among the institution's users even if it is inadequate to totally control their security behaviour. Where ISA programmes are in place, they need to know if they are sufficient, all of which is important, because many studies indicate that peoples' attitudes and their lack of security awareness are the most notable contributors to security incidents (Furnell, 2007).

Notably, as stated by Appleyard (2005), information is not the same in terms of its value to each institution, or in the threats that it is subject to. Institutions must implement regulations and rules of security to prevent illegal access to their information system assets. Technology alone is unable to assure a safe environment for information so the human aspect of IS must be considered along with the technological aspects (Sohrabi Safa, Von Solms and Furnell, 2016)

Organisations need to understand that any type of attack could result in a loss of their assets and reputation. Therefore, successful organisations are aware of the significance of protecting their information assets and of maintaining the highest level of IS (Waly, Tassabehji and Kamala, 2012). Organisations should not depend on technical approaches alone to protect their information assets, as there are still a number of security attacks that cannot be defeated by technical tools.

Regarding adherence to the ISP, motivational factors are as significant as sanctions or preventive elements in efficient information security management. Therefore, adopting a combination of the use of punishment and reward strategies can produce benefits from both approaches regardless of their individual drawbacks. Many studies

have found that adopting both procedures is fair to both the employers and employees, and could have a significant impact on compliance with the ISP (Bulgurcu, Cavusoglu and Benbasat, 2009, 2010; Chen, Ramamurthy and Wen, 2012).

Gaunt (1998) determines the success of an institution security policy by the clarity of the employee's rights and responsibilities. Biddinika et al. (2016) believe that 'readability is the starting point for the comprehension of a text' (p. 1351). In addition, Doherty and Fulford (2003) produce a set of factors including the complexity of the policy document that result in users ignoring security policies. Therefore, we can postulate that ease of reading is positively associated with security compliance intentions. Institutions should strive to achieve ISP compliance via comprehensibility. In light of changing circumstances and technological progress, ISPs should be regularly enhanced and updated to maintain fitness with the institution's vision and mission. The next chapter will explore what is meant by readability.

## **CHAPTER THREE: READABILITY**

### **3.1. Introduction**

Several factors enable compliance with regulations and rules of security and one of these is the comprehensibility of the information security policy (ISP) itself. In part, this can be estimated by applying a readability formula to the text of ISPs.

Procedures manuals provide instructions, directions, plans, and work routines. It is essentially kind of an implementation set of instructions of how policies would be implemented. In the context of information security, policies such kind a high level description of what needs to be achieved and the procedure is the set of instructions that given to people in order to make certain that they carry out this policy. On the context of this thesis, the term procedure is being used to explain the methodological steps that I have used to evaluate the readability in ISPs.

This chapter contains a detailed explanation of the literature that addresses readability and considers factors that can affect readability. Here, we present the development of readability measurements by introducing the most common traditional readability measurements and recent statistical approaches, discussing each in detail, as well as comparing between these two approaches. The chapter further reviews the criticisms and limitations of readability formulae.

### **3.2. Defining Readability**

There are several definitions that have been proposed for readability, for example, Klare describes readability as ‘the ease of understanding or comprehension due to style of writing’ (Klare, 1963, cited by Anagnostou and Weir, 2006). This explanation only covers the writing style, regardless of other factors such as the text’s format, coherence, and organisation (DuBay, 2004). McLaughlin considers more factors than DuBay, especially in respect of reader characteristics, like motivation, reading skill, and appropriate knowledge. He represents readability as ‘the degree to which a given class of people find certain material compelling and comprehensible’ (McLaughlin, 1969, cited by DuBay, 2004).

Many readability metrics are commonly based on quantifiable textual aspects such as length of words, the length of sentences, and a number of syllables or differences between these constructs (Weir & Anagnostou, 2008). However, according to Gray and Leary (1935), there are more than 220 factors that can affect readability. They classified factors into four groups (Content, Style, Format, and Features of Organisation). DuBay (2004, p. 18) highlights 17 factors out of 64 countable factors (suggested by Gray and Leary, 1935) as being the most significant factors, as this is shown in Table 3.1.

	<b>Factors Affect Readability</b>
1.	Average sentence length
2.	Percentage of “easy” words
3.	Number of words, not known to 90% of sixth-grade students
4.	Number of “easy” words
5.	Number of different “hard” words
6.	Minimum syllabic sentence length
7.	Number of explicit sentences
8.	Number of first, second, and third-person pronouns
9.	Maximum syllabic sentence length
10.	Average sentence length in syllables
11.	Percentage of monosyllables
12.	Number of sentences per paragraph
13.	Percentage of different words not known to 90% of sixth-grade students
14.	Number of simple sentences
15.	Percentage of different words
16.	Percentage of polysyllables
17.	Number of prepositional phrases

Table 3. 1: Factors that can affect readability

Because of the wide range of possible factors, it might be that certain factors are more appropriate in some circumstances than other. However, in reality, this may be problematic if some factors prove intractable to easy measurement. There are more

than 220 factors that can affect readability, e.g. reader's abilities cannot be measured mathematically. Readability metrics usually return an approximation of a text's difficulty. This is often expressed as a grade level, i.e., the years of educational study required to be capable of understanding the text (Anagnostou and Weir, 2006).

### **3.3. How to Evaluate Readability Measurement?**

There are over 200 metrics for measuring readability. A significant number of the assumptions made about readability are associated with readability formulae, since such formulae were devised (since 1948) for the purpose of evaluating the ease of reading for texts. There are many existence readability metrics that are beneficial, practical and objective predictors of text difficulty. In addition, there are widely used readability formulae due to their simplicity of operation and easy to calculate. Nevertheless, there are a number of metrics for measuring readability have been criticised attributable to their limitations.

#### **3.3.1. Standard Approach to Readability Measurements**

##### **3.3.1.1 Traditional Readability Formulae**

There are many traditional readability formulae for assessing the ease of reading for words in a piece of content. These metrics usually depend heavily for measurement on the shallow features of texts, such as counting syllables, characters per word, words, and sentences (Feng, Elhadad and Huenerfauth, 2009; Bailin and Grafstein, 2016). The readability formulae vary from one to another in the number, weights and type of factors adopted. In other words, the difference between readability formulae is made by their type of variables. The formulae depend on variables and specific weights for the variables. The calculation is based on the variables, which those variables are considered to contribute to making a text easy or difficult to read, whereas, specific weights for the variables show their relative significance (Bailin and Grafstein, 2016).

Several classic readability formulae utilise straightforward linear functions with two or three language features to model the readability of a given content. The features studied usually address two aspects: lexical and syntactic. Lexical features, aimed to measure the difficulty of vocabulary, often look at three predictor variables: the number of syllables a word includes, the number of characters a word includes, and



unfamiliar/difficult words. Vocabulary with more syllables or characters is counted to be harder. Frequently utilised words are supposed to be more readily understood than those that are less frequently encountered. In term of syntactic features, they are more limited than lexical features. Syntactic features, intended to measure the average sentence length (ASL) in words (Feng, Elhadad and Huenerfauth, 2009).

These characteristics can be seen in several common traditional formulae. For instance, the commonly used Flesch Reading Ease (FRE) and the Flesch-Kincaid grade level (FKGL) formulae count ASL and average syllables per word to return grade level scores (Anagnostou and Weir, 2006; Mesmer, 2008; Ammann and Sowa, 2013; Bailin and Grafstein, 2016). Both of the Gunning Fog Index readability formula (FOG) (Anagnostou and Weir, 2006; Sumeeth, Singh and Miller, 2010; Janan and David, 2012; Zhou, Jeong and Green, 2017) and the SMOG metric (Mc Laughlin, 1969; Anagnostou and Weir, 2006; Sumeeth, Singh and Miller, 2010; Bailin and Grafstein, 2016; Zhou, Jeong and Green, 2017) count ASL and the polysyllabic words to determine the grade level of a text. The Automated Readability Index (Senter and Smith, 1967; Zhou, Jeong and Green, 2017) and Coleman-Liau index (Wilson, Rosenberg and Hyatt, 1997; Zhou, Jeong and Green, 2017) are unlike other traditional readability metrics as they are based on characters instead of syllables per word, along with sentence length in words, to predict grade level. Different from the syllabic approach, the Spache and the “new” Dale-Chall formula made an advance in measuring lexical difficulty by utilising a list of familiar words. The Spache metric is similar to the “new” Dale-Chall metric in a sense that it counts two variables, ASL and number of unfamiliar words. However, the Spache formula was designed for measuring texts that are for children up to fourth grade. For older children, it is recommended to use the new Dale–Chall index (Bruce, Rubin and Starr, 1981; Redish, 2000; Anagnostou and Weir, 2006; Mesmer, 2008; Sumeeth, Singh and Miller, 2010; Benjamin, 2012; Janan and David, 2012; Bailin and Grafstein, 2016; Child, 2017; Crossley *et al.*, 2017; Readabilityformulas.com, 2017). To make the point clearer, Table 3.2 clarifies which of these predictor variables are utilised by each formula.

Formula	Predictor Variables				
	Number of Syllables	Number of Characters	Unfamiliar / Difficult words	Polysyllabic words	Sentence Length
FRE	✓				✓
FKGL	✓				✓
FOG				✓	✓
Spache			✓		✓
Dale-Chall			✓		✓
SMOG				✓	✓
Coleman-Liau		✓			✓
Automated Readability		✓			✓

Table 3. 2: Predictor variables used in the eight readability metrics

It is obvious from Table 3.2 that each of the introduced readability formulae use two factors to determine the grade level of a text and they share a common factor, which is the ASL variable.

With regard to the weights of factors used by each traditional readability metric, the readability formulae that have been commonly used as methods of assessing textual difficulty are illustrated below.

- Flesch Reading Ease:

$$206.836 - 1.015 \left( \frac{\text{total words}}{\text{total sentences}} \right) - 84.6 \left( \frac{\text{total syllables}}{\text{total words}} \right)$$

- Flesch-Kincaid Grade Level:

$$0.39 \left( \frac{\text{total words}}{\text{total sentences}} \right) + 11.8 \left( \frac{\text{total syllables}}{\text{total words}} \right) - 15.59$$

- Gunning FOG Index:

$$0.4 \left[ \left( \frac{\text{words}}{\text{sentence}} \right) + 100 \left( \frac{\text{complex words}}{\text{words}} \right) \right]$$

Where “complex words” are defined as words with  $3 \geq$  syllables

➤ Spache Index:

The original formula was:

$$\text{Grade} = (0.141 \times \text{average sentence length}) \\ + (0.086 \times \text{percentage of difficult words}) + 0.839$$

The revised formula (1978) is:

$$\text{Grade} = (0.121 \times \text{average sentence length}) \\ + (0.082 \times \text{percentage of difficult words}) + 0.659$$

➤ Dale-Chall Index:

$$\text{Grade} = 0.1579 \left( \frac{\text{difficult words}}{\text{total words}} \times 100 \right) + 0.0496 \left( \frac{\text{total words}}{\text{total sentences}} \right)$$

➤ SMOG Index:

$$1.0430 \sqrt{\text{number of polysyllables} \times \left( \frac{30}{\text{number of sentences}} \right)} + 3.1291$$

➤ Coleman-Liau Index:

$$\text{Grade} = (0.0588 \times \text{average number of letters per 100 words}) \\ - (0.296 \times \text{average number of sentences per 100 words}) \\ - 15.8$$

➤ Automated Readability Index:

$$4.71 \left( \frac{\text{total number of characters}}{\text{total number of words}} \right) + 0.5 \left( \frac{\text{total number of words}}{\text{total number of sentences}} \right) \\ - 21.43$$

### **Flesch Reading Ease Formula and Flesch-Kincaid Formula**

The FRE formula is probably the earliest readability metric and one of the most commonly used. This was published in 1948 by Rudolph Flesch and is based on the number of syllables and the number of sentences for each 100-word block of text

(DuBay, 2004). The results of this formula are calculated on a scale of 1 to 100, with less than 30 being text that is very complicated to understand and with greater than 90 being text that is very easy to understand (see Table A3.2, Appendix A) (Ammann and Sowa, 2013).

A modified version of the FRE formula called FKGL formula, was developed later (1975) by J. Peter Kincaid and other readability scientists assigned by the United States Navy. Both formula versions are based on the same core measures (word length and sentence length); however, they have different weighting factors that result in different readability scores (Anagnostou and Weir, 2006; Mesmer, 2008; Ammann and Sowa, 2013; Bailin and Grafstein, 2016).

Despite its popularity, in terms of common usage, the FRE has several recognised weaknesses. A principal concern is the sole reliance on ‘internal’ characteristics of the considered texts. Nevertheless, Crossley et al. (2017) state that the FRE formula created higher inter-correlations compared with the Automated Readability Index and the FOG index (op. cit.).

### **Gunning Fog Index**

This readability formula is another well-known traditional readability metric, developed in 1952 by Robert Gunning, and it is similar to FKGL except that it counts the percentage of words with more than two syllables instead of counting all of the syllables in a word. In addition, the FOG index was created specifically for adult level materials and became popular because it is user-friendly. The FOG index maps directly to educational grade-level (Anagnostou and Weir, 2006; Sumeeth, Singh and Miller, 2010; Janan and David, 2012; Zhou, Jeong and Green, 2017).

This formula has some limitations due to its function of counting the percentage of words with more than two syllables. Not all words with three or more syllables can be considered as difficult. There are also words with less than three syllables that would be considered difficult, especially if rarely used by most people. For example, the word “January” has four syllables and is not thought to be a difficult word (op. cit.).

## **Spache Formula**

The Spache formula is another traditional readability measure, developed in 1953 by George Spache. To revise the formula, another version was published in 1978. This measure uses two variables, sentence length and number of unfamiliar words (n= 769) (Bruce, Rubin and Starr, 1981; Janan and David, 2012; Child, 2017; Readabilityformulas.com, 2017).

The Spache readability index is similar to the Dale–Chall readability index, and utilises a list of familiar words. The Spache metric performs best on text documents that are intended for children up to fourth grade. For older children, it is recommended to use the new Dale–Chall index, as it is more appropriate than the Spache readability index (op. cit.).

Despite its popularity, in terms of common usage for primary school students, the Spache index has been criticised for its poor reliability. Researchers argue that the Spache formula’s set of familiar words have some discrepancies (Edwards and Gibbon, 1973; Janan and David, 2012).

## **Dale–Chall Readability Formula**

The original Dale–Chall index is another commonly used readability measure, created in 1948 by Edgar Dale and Jeanne Chall. This formula is similar to the Flesch reading formula in a sense that it uses two variables, ASL and a percentage of difficult words. However, it was designed to overcome some flaws of the FRE formula by assessing word complexity based on a large list of frequent words (n = 3000), at least 80 percent of which are familiar to fourth-grade students (Redish, 2000; Anagnostou and Weir, 2006; Mesmer, 2008; Sumeeth, Singh and Miller, 2010; Benjamin, 2012; Bailin and Grafstein, 2016; Crossley *et al.*, 2017).

The formula has been improved by including an updated list of familiar words in 1995 and called the “new” Dale-Chall formula. The “new” Dale-Chall metric has considered many readability research results in the almost half century since the original development and obviously has better predictive power than the initial formula (op. cit.).

In spite of its popularity, with regard to common usage, the Dale–Chall Readability formula has various criticisms as the formula’s output depends heavily on the frequent words list. First, vocabulary can change quite rapidly and different social and ethnic groups of people have different core vocabularies. Second, neither version of the Dale-Chall formula was created for measuring technical materials so it is questionable whether this formula gives an accurate score for technical and legal documents. In addition, the developers of the Dale-Chall metric did not aim to provide any theoretical insight into readability (Redish, 2000; Bailin and Grafstein, 2016).

### **SMOG Formula**

The SMOG index is one of the simplest and fastest metrics to compute. The formula was introduced in 1969 by G.H. McLaughlin. The idea of this formula is based on that semantic and syntactic difficulty predictors should be multiplied rather than added. The SMOG metric counts a number of sentences (at least 30) and in those sentences, counts the polysyllabic words (three or more syllables). The SMOG index returns grade level scores (Mc Laughlin, 1969; Anagnostou and Weir, 2006; Sumeeth, Singh and Miller, 2010; Bailin and Grafstein, 2016; Zhou, Jeong and Green, 2017).

Despite its simplicity, the SMOG formula has been criticised. In terms of accuracy, the SMOG metric to some extent is less accurate than other classic readability formulae as the formula publisher stated that his metric would predict grade level within 1.5 grades 68% of the time (Bailin and Grafstein, 2016; Zhou, Jeong and Green, 2017). Similarly, this formula and the FOG index based their count on the number of polysyllables. Not all words with three or more syllables can be deemed as difficult. Furthermore, the main focus of this formula, which is the sole reliance on ‘internal’ characteristics of the considered texts, is not that different from the other traditional formulae we have mentioned so far.

### **Coleman-Liau Formula**

The Coleman-Liau index is less commonly used than the other mentioned earlier formulae. Devised in 1975 by Meri Coleman and T. L. Liao, this formula is unlike other traditional readability metrics as it based on characters instead of syllables per word along with sentence length in words to determine grade level. The opinion of the formula’s developers is that a readability formula lacked accuracy if its techniques

relied on the number of syllables per word (Wilson, Rosenberg and Hyatt, 1997; Wilson and Wauson, 2010; Zhou, Jeong and Green, 2017).

The Coleman-Liau formula maps directly to the grade-level required to understand the text. It is worth mentioning that some researchers have noted that the Coleman-Liau index generates scores lower than the FRE formula when applied to technical documentation (op. cit.).

Although Meri Coleman and T. L. Liau argued that their formula is more accurate than other traditional readability formulae, there is a clear weakness in all traditional readability formulae as mentioned earlier, since their account relies on ‘internal’ characteristics of the considered texts.

### **Automated Readability Index**

This formula is another traditional readability formula for testing English language texts, published in 1967 by Smith and Senter, in an attempt to evaluate the readability of written documents used by the US Air Force. This formula is similar to the Coleman-Liau formula, in term of core measurement. They use same variables (characters instead of syllables per word and average number of sentence in words) and return grade level scores (Senter and Smith, 1967; Colmer, 2017; ReadabilityFormulas.com, 2017; Zhou, Jeong and Green, 2017).

The Automated Readability Index and other traditional readability formulae have been criticised as these formulae have flawed and their account reliance on ‘internal’ characteristics of the considered text documents (Anagnostou and Weir, 2006; Janan and David, 2012; Bailin and Grafstein, 2016; Zhou, Jeong and Green, 2017).

The characteristics of the classic readability formulae were outlined above and it is worth mentioning that, the readability metrics described above were used in analysing a set of eight sample policies of our study.

#### **3.3.1.2 Coh-Metrix**

Coh-Metrix is a web-based software tool was developed at the University of Memphis (in the United States) as a computational tool that able to calculate the coherence of texts on a wide range of measures as well as include indices of the

linguistic and discourse representations of a text document. In general, the term Coh-Matrix refers to a computer program that is able to examine text difficulty on more than 50 types of cohesion relations and over than 100 measures of language, text and readability as well as features of cohesion that signify a text's level of coherence (Graesser *et al.*, 2004, 2014a). Cohesion means 'surface indicators of how sentences are related to one another in a text (a text construct)' (Benjamin, 2012, p. 73). One of the powerful features of Coh-Matrix software is using latent semantic analysis (LSA) which is defined by Benjamin (2012, p. 70) as 'an automated tool that represents text content as a vector in semantic space'. This tool assures as a method for measuring the readability of texts particularly in relationship with reader background knowledge. A practical study conducted by Foltz *et al.* (1998) revealed that LSA significantly better than basic word overlap measure. Coh-Matrix version 3.0 contains 106 indices of cohesion, language, and three of them are used to assess readability or text difficulty. It included FRE, FKGL and Coh-Matrix. This software has its own formula, which is specifically developed to predict text difficulty for second language readers (called Coh-Matrix L2 Readability) (Graesser *et al.*, 2014b).

Coh-Matrix readability measure uses three variables: word overlap, sentence syntactic similarity, and word frequency. The syntactic similarity variable is measured via number of words before the main verb. This automated tool, Coh-Matrix, has a distinguishable world overlap variable that the function of this variable can analyse how frequently two sentences share regular arguments (i.e. nouns, pronouns, and noun phrases). The Coh-Matrix formula utilises extrinsic factor by calculating word frequency information over CELEX frequency scores. The frequencies of CELEX are taken from the early 1991 version of the COBUILD corpus (Crossley *et al.*, 2007). The formula itself was based on the subset of Bormuth's (1971) corpus of 32 academic texts. The full detail of the Coh-Matrix equation cannot be found, however, as reported in McNamara *et al.* book (McNamara *et al.*, 2014) the Coh-Matrix L2 Readability formula is that:

### **3.3.1.3 Strathclyde Readability Formula**

Of the available formulae for evaluating the readability of text, there is a formula called the Strathclyde Readability Measure (SRM) - developed at our institution as a



‘new generation’ of readability metric- which differs in approach from most other readability measures. Traditional readability metrics often depend for measurement on counting syllables, characters per word, words, and sentences only (intrinsic factors), but SRM differs in taking into account the frequency of occurrence of words, relative to the British National Corpus (extrinsic factor) (Anagnostou and Weir, 2006). Instead of average sentence length (ASL), the SRM employs a constant based on ASL, in order to obtain scores that are closely associated when texts are similar in difficulty but different in ASL. The SRM provides two versions. SRM1 is designed for texts that have more than 150 words while SRM2 is suitable for texts with less than 150 words (op. cit.).

Two of the following equations (Weir and Ritchie, 2006, p. 4) give the SRM scores:

$$SRM1 = \{\log \times (AWF \times 2) \times K\} - 80$$

Where:

- ❖ AWF = the average word frequency, only calculating words with a frequency not more than 100,000.
- ❖ K = a constant depends on the ASL
  - 15: if the ASL  $\geq 17$  and  $< 25$ , or the ASL  $< 17$  and the AWF  $> 95000$ .
  - 13: if the ASL  $< 17$  or  $\geq 25$ .

$$SRM2 = 100 - \{[(cc \times 100) \times K] \times 3\}$$

Where:

- ❖ cc = number of difficult words divided by the sum of words in the text. Where a difficult word is a word with a frequency less than 100.
- ❖ K = a constant depends on the ASL
  - 3: if the ASL is  $\geq 17$  and  $< 25$ , or the ASL is  $< 17$  and the AWF is  $> 95000$ .
  - 5: otherwise.

The effects of this formula are measured on a 100 scale, with less than 30 reflecting complicated text and greater than 80 reflecting easy to read a text. Generally, the range

of SRM differs from the FRE formula in estimating readability (see Table A3.1, Appendix A).

### **3.3.2. Another Approach to Readability**

#### **Cloze Test for Comprehension Measurement**

The previous approaches use formulae to assess readability whereas the Cloze test measures the degree of user comprehension and does not offer a measure of the document comprehensibility.

Many test types seek to measure understanding. The Cloze procedure is a method for gauging reading comprehension. The appearance of Cloze test as a tool for measuring readability in 1953 stimulated the development of new criterion passages, and this type of measurements ensures that judgment of readability is not determined solely on a mechanical basis (DuBay, 2004). Richards and Schmidt (2015) explain that reading is the process of perceiving a text for the purpose of understanding its content. The understanding that results is described as reading comprehension.

The Cloze test is well known, especially for testing language abilities. This test includes a text with a number of deleted or removed words from a reading passage, where the test taker is required to fill in the missing words. The test designer usually chooses one of two techniques to create the blanks. The first is called rational deletion (rational Cloze), as the name suggests where the test creator decides which words are deleted based on some rational or criterion principle. The second technique is known as fixed ratio deletion or  $n^{\text{th}}$  word deletion, where every  $n^{\text{th}}$  word is removed systematically. Thus, the test taker is required to construct meaning from the passage by identifying the missing words (Kobayashi, 2009; Richards and Schmidt, 2015).

A quote passage from NIST special publication (Wilson and Hash, 2003, p. 18) is provided below to show an example of Cloze test with rational deletion and fixed-ratio deletion.

### Example of Rational Deletion

An awareness program ..... begin with an effort that can be ..... and implemented in ..... ways and is aimed at all levels of the organization including senior and executive ..... . The effectiveness of this ..... will usually determine the effectiveness of the ..... and ..... program. This is also true for a successful IT security program.

### Example of Fixed-ratio Deletion (6<sup>th</sup> Word Deletion)

An awareness program should begin ..... an effort that can be ..... and implemented in various ways ..... is aimed at all levels ..... the organization including senior and ..... managers. The effectiveness of this effort ..... usually determine the effectiveness of ..... awareness and training program. This ..... also true for a successful ..... security program.

Figure 3. 1: Cloze test's types example

For the example of Cloze test with rational deletion (see Figure 3.1), the words chosen to be 'missing words' were considered significant in their contribution to the meaning of the phrase in which they appeared. Whereas the second example of cloze test with fixed-ratio deletion, the words chosen to be 'missing words' were decided to be the sixth word regardless of the word type.

There are two types of scoring Cloze test performance; the first type is called the exact word method. The score will be counted if only the test taker writes the exact name that was deleted from the original text. The second one is known as acceptable word method (or the appropriate word method, the acceptable alternative method, or the contextually appropriate method). In this type, the score will be counted when responses are suitable or acceptable in the context. To be precise, the acceptable word method split into two categories, which are syntactically and semantically acceptable word scoring or semantically only acceptable word scoring method (Kobayashi, 2009; Richards and Schmidt, 2015).

Singer and Donlan (1982), and Biddinika et al. (2016) confirm that readability can be measured by using two approaches which are the readability formula and reader responses. Reader responses are a group of tests on the ability toward read a text such as Cloze tests. There are many authors suggesting using the Cloze procedure as a method of testing the readability of a text for reading comprehension (Taylor, 1953; Bormuth, 1967; Rankin and Culhane, 1969; Guillemette, 1989; Kobayashi, 2012; Richards and Schmidt, 2015).

As noted, several formulae have been explored in the literature as a basis for gauging the readability of written material, but invariably depend upon syntactic variables (Klare, 1985). Traditional readability metrics have several limitations, including 1- they generally focus on purely internal characteristics of the considered texts and ignore the likely familiarity or unfamiliarity of the terms found therein (Campbell and Weir, 2006), 2- They are generally insensitive to whether the texts are meaningful or senseless, 3- there is variation in the results of readability metrics for the same content, 4- readability formulae assume that people are similar in characteristics, maturity and skills (Anagnostou and Weir, 2006).

### **3.4. Comparison of Approaches to Readability**

Despite the usefulness of using readability formulae to easily and quickly evaluate written text difficulty for readers, it is worth to keep in mind that they have some shortcomings. To date, several scholars in the field of readability research criticised the use of readability formulae due to its flaws. The main disadvantages of traditional readability formulae that have been recognised in the relevant literature can be listed as follows:

- **Cannot suggest how to improve the comprehensibility of the evaluated document:** traditional readability formulae, Strathclyde formula and Coh-Metrix readability scores tell you how hard the text to read and the suitable group of age for reading the evaluated piece of text, but they do not offer suggestions for improving the comprehensibility of a text document (Kobayashi, 2009).
- **Fail to discover incomprehensibility of expression:** readability scores of the three introduced approaches, traditional readability formulae, Strathclyde

formula and Coh-Metrix tool, remain the same even for scrambled sentences that have no meaning (Anagnostou and Weir, 2006; Benjamin, 2012; Crossley *et al.*, 2017).

- **Fail to accurately indicate difficult words in a text:** traditional formulae presume that texts with short words and sentences, and words with less syllables are easier to read and comprehend by the average people. This assumption is not always accurate, there are many instances of texts with short but complex word or words with more syllables but is not thought to be a difficult word (Anagnostou and Weir, 2006; Sumeeth, Singh and Miller, 2010; Janan and David, 2012; Zhou, Jeong and Green, 2017). However, Strathclyde formula overcome this shortcoming by utilising corpus analysis to find words that appear less frequently than others from a frequency list generated from the target text. To put it another way, Strathclyde formula estimates text complexity not only on counting word length but also on frequency of occurrence relative to the British National Corpus. In similar fashion, Coh-Metrix uses three variables: word overlap, sentence syntactic similarity, and word frequency to overcome the traditional formulae deficiencies.
- **Assuming readers are alike:** traditional and Strathclyde readability metrics cannot make distinctions based on reader's characteristics. Which is no consideration, be taken by readability index, of maturity, ability and the first language of the targeted readers (Redish, 2000; Anagnostou and Weir, 2006). On the other hand, Coh-Metrix has a feature of measuring the readability of texts particularly in relationship with reader background knowledge (Benjamin, 2012).
- **Only able to measure what can be counted:** readability metrics can measure for example: sentence length, syllables per word or whether each word is on a list of acceptable words, but they cannot measure, the number of inferences demanded, interest level, rhetorical structure, difficulty of concepts, or dialect (Redish, 2000; Anagnostou and Weir, 2006; Mesmer, 2008; Janan and David, 2012). Furthermore, readability metrics gauged text difficulty and ignoring other important aspects such as the context of its use, the structure of the text,

enthusiasm, curiosity, competitiveness, value and reading purpose, etc. (Bruce, Rubin and Starr, 1981).

- **Only accept one language:** traditional readability formulae, Strathclyde formula and Coh-Metrix are English-based formulae so they can only measure a piece of English text.
- **Impaired by formatting or punctuation:** the three comprehension approaches to readability, Traditional readability formulae, Strathclyde formula and Coh-Metrix, do not work efficiently on a number of aspects. They do not provide accurate results if the text on forms, web pages, or text documents with many punctuation marks, bulleted lists, headings, equations, codes, tables and figures (Redish, 2000; Si and Callan, 2001; Collins-Thompson and Callan, 2004; Schwarm and Ostendorf, 2005; Benjamin, 2012; Zhou, Jeong and Green, 2017) .
- **Depending on intrinsic syntactic features:** many classic readability metrics reliance on only internal characteristics of the considered text documents and ignore the likely familiarity or unfamiliarity of the terms found therein. This could lead a formula to present an inaccurate and misleading score, since the traditional readability metrics utilise heuristic measures based on internal qualities of a document. This is often as a result of the reader's own understanding of a given text document is being either inaccurately measured or occasionally being entirely overlooked. Weir and Ritchie (2006) emphasise that most of the classic readability formulae rely on intrinsic syntactic features, which is an inherent limitation of most traditional readability measures, as they have no reference to the probability that readers will be familiar with the constituent words and expressions. In contrast, Strathclyde formula and Coh-Metrix reliance on internal and external characteristics of the measured texts.
- **Producing disparate reading levels for the same piece of text:** the difference in readability scores between readability metrics are the result of the different variables and different criterion scores adopted by different readability metrics. Previous published studies have highlighted this shortcoming (Klare, 2000; Zhou, Jeong and Green, 2017).

In term of Cloze test for comprehension, we can say that the prominence of the Cloze test can be credited to its simplicity of construction and administration if the exact word method is utilised. Moreover, its reputed effectiveness as a measure of worldwide language competence. However, the test should be carefully designed in order to achieve text validity and to assess readability rather than textual redundancy (Kobayashi, 2009). It is worth to mention that there is a factor that Cloze test designer should carefully be aware of which is the effects of basic text display variables on readability including font size, border weight and colour (Bernard *et al.*, 2003; Darroch *et al.*, 2005; Slattery and Rayner, 2009).

Cloze procedure could require a lot of time and effort when the acceptable word method is used as scoring method. The test marker will spend a lot of time on determining whether responses are suitable or acceptable in the context or not. Furthermore, Cloze test requires a lot of participants to enroll for results' reliability and validity, conversely, readability formula is a software tool not require human participation. Moreover, not all Cloze procedure measure exactly the same abilities, the result of Cloze test could vary based on A) the chosen technique (rational deletion or fixed ratio deletion), B) the scoring method (exact word method or acceptable word method), C) the starting point for deletion, and D) the deletion rate (Alderson, 1983; Fuchs, Fuchs and Maxwell, 1988; Kobayashi, 2009).

This study has found that the automated tool Coh-Metrix, is distinguished by its world overlap variable from Strathclyde readability tool, as the function of this factor is to analyse how frequently two sentences share regular arguments (i.e. nouns, pronouns, and noun phrases). In addition, The Coh-Metrix tool has a feature of measuring the readability of texts particularly in relationship with reader background knowledge. On the other hand, the Coh-Metrix formula is similar to Strathclyde formula in term of utilising extrinsic factor by calculating word frequency information over CELEX frequency scores.

It is claimed that the Coh-Metrix index measures text challenges at the sentence and the word level, as well as possessing complex features such as the cohesion between sentences in a text. Nevertheless, the feature of measuring cohesion has not escaped criticism from Mr Benjamin (2012) as he said that the advanced Coh-Metrix

variables did not show any difference in performance from the traditional variables. One of the limitations with Coh-Metrix is that this automated tool is based on only 32 academic reading texts of Bormuth's (1971) corpus and this amount of academic texts not certainly representative of all type of texts. Crossley et al. (2007) commented on the latest point by arguing that the Coh-Metrix L2 readability may not be generalizable unless it uses larger corpora.

The study of Coh-Metrix would have been more useful and relevant if it had studies on native speaker participants along with second language participants, a formula to predict text difficulty for first language readers, and explored a wide range of the most common readability formulae. Nonetheless, this software programme is easy to use, accessible via a straightforward web interface, beneficial for measuring readability of second language readers, has many indices enabling it to analyse potential comprehensibility more than traditional readability metrics, covering the limitation of readability formulae of not measuring the cohesion, and has many functions and features that make it stands out as a useful tool.

All in all, despite readability formulae' shortcomings, they are still in widespread use today because they fill a need. Given such a level of critique, it may lead to wonder whether there is a usefulness of adopting readability formulae to estimate the text difficulty. Although readability metrics have their limitations, they are beneficial, convenient, easy-to-use, practical and objective predictors of the difficulty of a text nevertheless should be applied accurately, along with other approaches to readability measurement and with cognisance of their shortcomings (see Table 3.3). For such reasons, adopting a Cloze test - a human-based comprehension test - ensures that judgment of readability is not determined solely on a mechanical basis. The Cloze test measures the degree of user comprehension rather than the document comprehensibility. In addition, Kobayashi (2009) emphasises that there is a high correlation between readability metric scores and comprehension test results. Bormuth (1966) believed that Cloze procedure is suitable for measuring not only the difficulty of the whole text but also the difficulty of individual words, phrases, and clauses. Part of our objective is to compare the use of a software readability approach with human



comprehension tests as a basis for insight on the readability of ISP documents. This should shed light on whether readability may influence the understanding of ISPs.

<b>Statements</b>	<b>Traditional Formulae</b>	<b>Strathclyde Formula</b>	<b>Coh-Metrix tool</b>
Cannot suggest how to improve the comprehensibility of the evaluated document	Yes	Yes	Yes
Fail to discover incomprehensibility of expression	Yes	Yes	Yes
Fail to accurately indicate difficult words in a text	Yes	No	No
Assuming readers are alike	Yes	Yes	No
Only able to measure what can be counted	Yes	Yes	Yes
Formulae can only accept one language	Yes	Yes	Yes
Impaired by formatting or punctuation	Yes	Yes	Yes
Depending on the shallow features of texts	Yes	No	No

Table 3. 3: Comparison between readability automated tools

### 3.5. Summary

This chapter provides the review of the literature on readability. Therefore, one of the methods of calculating ease of understanding is by using readability measurements. My literature review noted that a significant number of the assumptions made about readability are consequently connected with readability formulae. My literature review also noted that the readability formulae provide useful insight, but it cannot be used alone to estimate the readability of a text as equation scores of readability formula should not be treated as precise estimates. The referenced literature has outlined traditional readability formulae, comprehension test and the recent statistical approaches. Additionally, this chapter presented in detail the readability formulae, which will be used in our study as well as the limitations of readability formulae. The following chapter will clarify the different methods that have been adopted to accomplish the research objectives.

## **CHAPTER FOUR: RESEARCH METHODOLOGY**

### **4.1. Introduction**

This chapter presents a comprehensive clarification and justification of the adopted research process design. This includes research philosophy, design, and methodology. It discusses the philosophical worldview or the research paradigm, which leads the study process as a whole, and hence affects the select of methodology and methods to conduct such research.

This chapter also highlights and justifies the undertaken approach for this research as well as the employed qualitative, quantitative, and mixed method strategy. The data has been collected from a wide range of participants by experts' insights, focus groups, and surveys. At the end of the chapter, the importance of ethical assurance was emphasised.

### **4.2. Understanding Methodology**

Pickard described methodology as 'perspective, the angle the researcher wishes to take on the question being asked' (2013, p. 2). If the question starts with 'when', 'how many', or 'how often', then it would be called a quantitative angle. In contrast, it would be called a qualitative angle if the question starts with 'why', 'how', or 'how do they feel about it' (Pickard, 2013). In addition, methodology refers to a group of thoughts regarding the connection between phenomena or it is the theory of how scientists obtain knowledge in study contexts and why. Researchers can answer the 'why' question by explaining the purposes of adopting particular strategy and methods to construct, collect and develop specific types of knowledge regarding educational phenomena (Briggs, Morrison and Coleman, 2012). Collis and Hussey indicated methodology as 'an approach to the process of research, encompassing a body of methods' (2013, p. 55). The method is a procedure to obtain and analyse data (op. cit.).

Jonker and Pennink defined methodology in general as an explicit approach of structuring one's thinking and actions with regard to research (2010). Saunders, Lewis and Thornhill indicated methodology as 'the theory of how research should be

undertaken' (2009, p. 34). Whereas, Strauss and Corbin defined qualitative methodology as 'theoretical perspective into the social world. It is a way of thinking about and studying social reality' (1998, p. 3). whilst, Jayaratna defined quantitative methodology as 'an explicit way of structuring one's thinking and actions' (1994, p. 37). Therefore, it is strongly recommended for researchers to carefully comprehend all the assumptions underlying the research questions to employ the most suitable methodology.

### **4.3. Research Approaches**

A researcher should outline the research approach of his/her research. Highlighting the approach helps readers to know the study's strategies and procedures that span the stages from wide assumptions to specified methods of data collection, analysis and interpretation. The general choice includes which approach ought to be utilised to study a subject. Informing the selected (broad) research approach should be the philosophical assumptions the examiner delivers to the study, research designs and specific research methods of data collection, analysis, and interpretation (see Figure 4.1). It is worth to mention that the research problem, the researcher's personal experience and the study participants could also affect the choice of research approach (Creswell and Creswell, 2018).

Creswell and Creswell (2018), Creswell (2015), and Newman and Benz (1998) identified three research approaches, which are 1) qualitative, 2) quantitative, and 3) mixed methods. The three approaches are defined below, in order to underline their differences and what they represent.

Qualitative research is an approach to analysing and comprehending the meaning individuals or groups ascribe to a social or human problem. The term quantitative research applies to an approach for verifying objective theories by analysing the relationship between variables. Recently, there has been renewed interest in using mixed-methods designs by social and health sciences researchers. Mixed methods research can broadly be defined as an approach to investigation that includes the gathering of quantitative and qualitative data, integrating the two forms of data, and using distinct designs that might include philosophical assumptions and theoretical frameworks (Creswell and Creswell, 2018). The mixed methods process starts with

collecting, analysing, and combining, or integrating qualitative and quantitative data at some point in the research process within a single study in order to obtain a better understanding of the research problem (Teddlie and Tashakkori, 2010).

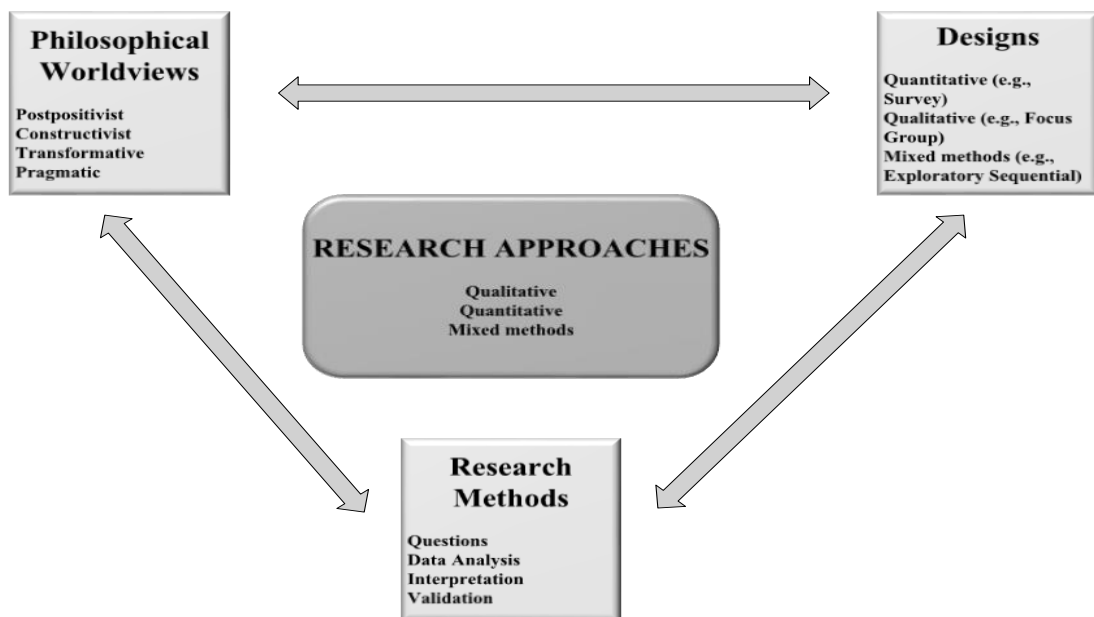


Figure 4. 1: A framework for research extracted from (Creswell and Creswell, 2018, p. 5)

Figure 4.1, represents the interaction between three components (Philosophical worldviews, Designs, and Research Methods). Researchers in planning procedure need to think carefully through the philosophical worldview assumptions that they deliver to the project, the research design that is associated with this philosophical worldview, and the specific research methods that turn the approach in the direction of practice (Creswell and Creswell, 2018).

#### 4.3.1. Philosophical Worldviews

Creswell and Creswell (2018) identified the major philosophical worldviews (paradigms) underlying different research approaches utilised in social researchers' work. Guba (1990, p. 17) characterise 'worldview' as 'a basic set of belief[s] that guide action'. Other scholars (Lincoln, Lynham and Guba, 2011; Briggs, Morrison and Coleman, 2012; Mertens, 2014; Bryman, 2015) have called worldviews paradigms which explain as an approach to research that delivers a uniting framework for

comprehending each of knowledge, truth, values and the nature of beings (Briggs, Morrison and Coleman, 2012).

Researchers are urged to make explicit the philosophical ideas they espouse when preparing their research proposal in order to express their reasons for selecting a quantitative, qualitative, or mixed methods approach for their research. There is an ongoing debate amongst researchers about the best approach to view and investigate social phenomena, which some scholars have called “the paradigm debate” (Savin-Baden and Major, 2013; Klenke, 2016). In the present research, we will outline four paradigms that are broadly discussed in the literature: a) post-positivism, b) constructivism, c) transformativism, and d) pragmatism. These philosophical worldviews reveal numerous core assumptions that are interrelated and mutually supportive of each other (Reber, 1995; Savin-Baden and Major, 2013; Lombardo, 2017). After we highlight these paradigms, a summary table (see Table 4.1) will be presented to show the major elements of each position.

### **Post-positivist Worldview**

The first group of researchers holds to the traditional form of research, which is the philosophical assumption of the post-positivist approach. Put simply, the positivists presume that there is a single reality, which can be measured and recognised, and thus this approach is typically seen more in quantitative research rather than qualitative research to measure this reality (Lincoln, Lynham and Guba, 2011). In contrast, the post-positivist term signifies the thinking after positivism, challenging the conventional notion of the supreme truth of knowledge and knowing that we cannot be totally positive about our claims of knowledge when considering the behaviour and actions of individuals (Phillips and Burbules, 2000; Creswell and Creswell, 2018).

Some scholars also call this the positivist, positivism, empirical science, or critical social theory (Briggs, Morrison and Coleman, 2012; Savin-Baden and Major, 2013; Bryman, 2015; Creswell and Creswell, 2018). Supporters of this approach utilise experimental and quantitative techniques to confirm or reject a given hypothesis for the purpose of generalisations through the deduction process. The main aim of the approach is to define any causal link between the study variables and pose this in terms of hypotheses (Phillips and Burbules, 2000; Creswell and Creswell, 2018).

## **Constructive Worldview**

Another group of researchers holds to the philosophical assumptions of the constructive approach. Social constructivists presume that there is no single reality or truth and consequently, reality requires to be interpreted, and thus this approach is more likely to employ qualitative methods to obtain those multiple realities (Lincoln, Lynham and Guba, 2011).

This paradigm is some scholars called it also the constructivism, social constructivism, or interpretivism (Briggs, Morrison and Coleman, 2012; Savin-Baden and Major, 2013; Bryman, 2015; Creswell and Creswell, 2018). In this model, knowledge and truth are developed rather than discovered. Supporters of this school of thought generally invent concepts, models and schemes in order to make sense or interpret the meanings other have about the world. Rather than beginning with a theory (as in the previous paradigm), inquires or inductively develop a theory or pattern of meanings (Schwandt, 2000; Savin-Baden and Major, 2013; Creswell and Creswell, 2018).

## **Transformative Worldview**

Another position on worldviews comes from the transformative approach. The reason behind this approach is that some people felt there are some aspects that are not covered by the post-positivist paradigm (Mertens, 2014). Additionally, the Creswell and Creswell (2018, p. 9) explain that ‘this position arose during the 1980s and 1990s from individuals who felt that the post-positivist assumptions imposed structural laws and theories that did not fit marginalized individuals in our society or issues of power and social justice, discrimination, and oppression that needed to be addressed.’

Researchers use this approach when their main research is about diverse groups, to show marginalised community voices, or they want to focus on inequities based on some specific perspective (e.g. gender, race, sexual orientation...etc.) that result in asymmetric power relationships. Supporters of this school of thought generally link political and social action to these inequities and they usually utilise mixed methods to outline the ecological complexity of a circumstance and to reach the voices of individuals who have previously been marginalised (Mertens, 2014; Creswell and Creswell, 2018; Jackson *et al.*, 2018).

## Pragmatist Worldview

Others hold a different worldview. Pragmatism is generally considered as the philosophical partner of the mixed methods approach. Pragmatists presume that reality is always renegotiated, debated, interpreted, and hence the best technique to utilise is the one that solves the issue (Lincoln, Lynham and Guba, 2011; Creswell and Creswell, 2018). Researchers when using this paradigm usually focus on the research problem and question and utilise all approaches available to comprehend the problem, in place of focusing on methods (Rossman and Wilson, 1985). Many scholars highlighting the importance of giving attention on research problem from the beginning by emphasising that researchers in social science should initially focus on the research problem and then utilising the pluralistic approaches to derive knowledge about the problem (Patton, 1990; Morgan, 2007; Teddlie and Tashakkori, 2010; Creswell and Creswell, 2018).

Some researchers agree that pragmatism provides a philosophical basis for research (Cherryholmes, 1992; Morgan, 2007; Savin-Baden and Major, 2013; Creswell and Creswell, 2018). Supporters of this school of thought believe that researchers have freedom of choice; therefore, they can select whatever methods, techniques, and research procedures suit their requirements and purposes. Pragmatism is not obligated to one system of philosophy and reality. This applies to mixed methods research in that inquires draw liberally from quantitative and qualitative assumptions when they involve in their study. Supporters of this school of thought see the world as multi-units. So also, mixed methods researchers consider several ways for collecting and analysing data instead of subscribing to only one approach. Truth is what works at the time. Therefore, the pragmatist researchers employ both quantitative and qualitative data as they work to deliver the best understanding of research problem. The pragmatist researchers require establishing a purpose for their mixing, a rationale for the reasons why data required to be mixed (quantitative and qualitative). Therefore, the pragmatic worldview allows researchers to adopt multiple methods, different assumptions and different forms of data collection and analysis.

<b>Post-positivism</b>	<b>Constructivism</b>
<ul style="list-style-type: none"> <li>• Determination</li> <li>• Reductionism</li> <li>• Empirical observation and measurement</li> <li>• Theory verification</li> </ul>	<ul style="list-style-type: none"> <li>• Understanding</li> <li>• Multiple participant meanings</li> <li>• Social and historical construction</li> <li>• Theory generation</li> </ul>
<b>Transformative</b>	<b>Pragmatism</b>
<ul style="list-style-type: none"> <li>• Political</li> <li>• Power and justice-oriented</li> <li>• Collaborative</li> <li>• Change-oriented</li> </ul>	<ul style="list-style-type: none"> <li>• Consequences of actions</li> <li>• Problem-centred</li> <li>• Pluralistic</li> <li>• Real-World practice oriented</li> </ul>

Table 4. 1: Four philosophical worldviews

#### 4.3.2. Research Designs

A research design (also called strategies of inquiry) indicates the plan that constitutes the research study. It is the methods by which the objectives or aims of the study are satisfied (Briggs, Morrison and Coleman, 2012; Denzin and Lincoln, 2018). Moreover, Creswell and Creswell (2018) state that the task of the researcher does not stand on selecting a qualitative, quantitative, or mixed methods study to conduct, but also to decide which type of study within these three options.

The research designs, that are available for social science researchers, can be listed as follows:

#### Quantitative Designs

In quantitative research, research designs associated with quantitative approach were those that invoked the post-positivist paradigm and that introduced mainly in psychology during the late 1990s and through the 20<sup>th</sup> century such as true experiments and the quasi-experiments (Creswell and Creswell, 2018).

Causal-comparative research is a nonexperimental form of research in which the researcher compares between no less than two groups in terms of a cause. The Correlational design is another type of nonexperimental quantitative design in which researchers utilise correlational statistics to describe and measure the degree or relationship between variables or sets of scores. The survey design is one more type of



non-experimental designs (Creswell, 2012). Currently, quantitative approaches have included complex experiments with several variables and treatments such as factorial designs and repeated measure designs.

### **Qualitative Designs**

During the 19<sup>th</sup> and into the 21<sup>st</sup> century, more types and numbers of approaches become noticeably visible. Creswell and Creswell (2018) assert that the historical origin for qualitative research originates from anthropology, sociology, the humanities, and evaluation.

There are various types of qualitative inquiry approaches that viable techniques to conduct qualitative studies. These include, but is not limited to, narrative research, phenomenological research, participatory action research, ground theory, discourse analysis, ethnography, and case study (Moustakas, 1994; Stake, 1995; Clandinin and Connelly, 2000; Cheek, 2004; Kemmis and McTaggart, 2005; Wolcott, 2008; Fetterman, 2010; Charmaz, 2014; Corbin and Strauss, 2015; Yin, 2017; Creswell and Poth, 2017; Creswell and Creswell, 2018).

### **Mixed Methods Designs**

Mixed methods include the mixing of qualitative and quantitative research and data in a single research study. Qualitative data usually contains open-ended responses without predetermined answers. In contrast, quantitative data tends to reflect closed responses, for example, as in surveys or psychological instruments (Creswell and Creswell, 2018). The rationale for ‘mixing’ both kinds of data within one single study is grounded in the fact that quantitative methods are effective for providing a solution to some questions and qualitative methods are applicable for others and due to the benefits of combining or integration of qualitative and quantitative methods, many scholars emphasise that both methods can be combined in one study (Pickard, 2013; Zikmund *et al.*, 2013; Bryman, 2015; Creswell and Creswell, 2018). Johnson and Turner (2003) presented the fundamental principle of mixed research and Johnson & Onwuegbuzie (2004, p. 18) explain it by saying that ‘researchers should collect multiple data using different strategies, approaches, and methods in such a way that the resulting mixture or combination is likely to result in complementary strengths and non-overlapping weaknesses.’

There are many designs of mixed methods but we will introduce the three core designs found in the social sciences as shown in Figure 4.2 (Creswell and Creswell, 2018). Firstly, convergent mixed methods, in which the investigator converges or merges both quantitative and qualitative data, in a single-phase approach, for the purpose of providing a comprehensive response to the research problem. The second approach is the explanatory sequential mixed methods. In this design, the researcher firstly conducts quantitative research, analyses the outcomes and subsequently expands on the outcomes to clarify them in further detail through qualitative research. This is regarded as explanatory since the preliminary quantitative data results are explained further through the qualitative data. This is regarded as sequential since the preliminary quantitative part is followed by the qualitative part. The third mixed method design is termed exploratory sequential mixed methods. This is the opposite sequence from the explanatory sequential design. In this approach, the investigator starts with a qualitative research stage and explores participants' views. Subsequently, the qualitative data are analysed in order to use the information to feed into the second stage, the quantitative research (Creswell, 2013). An overview of these strategies of inquiry is presented in Table 4.2.

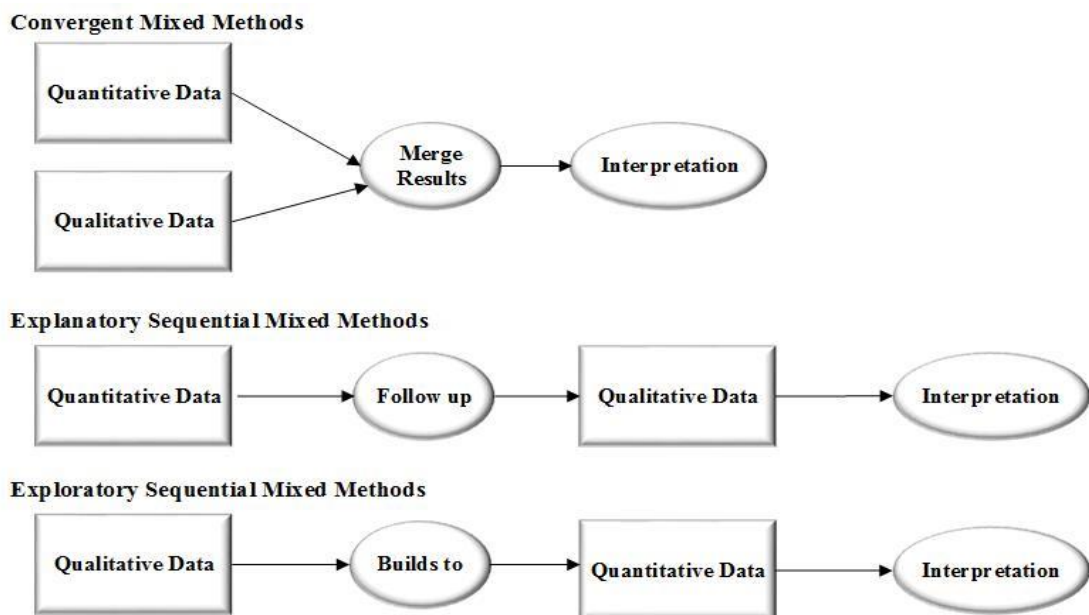


Figure 4. 2: Three mixed methods designs, extracted from (Creswell and Creswell, 2018, p. 218)

<b>Quantitative</b>	<b>Qualitative</b>	<b>Mixed Methods</b>
<ul style="list-style-type: none"> <li>• Experimental designs</li> <li>• Non-experimental designs</li> <li>• Longitudinal designs</li> </ul>	<ul style="list-style-type: none"> <li>• Narrative research</li> <li>• Phenomenology</li> <li>• Ground theory</li> <li>• Ethnographies</li> <li>• Case study</li> </ul>	<ul style="list-style-type: none"> <li>• Convergent</li> <li>• Explanatory sequential</li> <li>• Exploratory sequential</li> </ul>

Table 4. 2: Research designs' types

### 4.3.3. Research Methods

The third main component of the research framework is the precise research method that includes the procedures of data collection, analysis and interpretation that an investigator suggests for their studies. In some research types, quantitative and qualitative data are collected, analysed and interpreted. Creswell and Creswell (2018) introduce a wide range of possibilities of data collection which can be adopted by a researcher in their works as it is shown in Table 4.3.

Researchers have a variety of ways to collect data, either by using a tool, test, or by gathering information on a behavioural checklist. Moreover, collecting data may require an investigator to visit a research site and notice the behaviour of people without prearranged questions or organising an interview in which the participant can talk, without restrictions and the use of specific questions, about a topic. The kind of data analysed might be numeric information collected on scales of instruments or text information recording and reporting the voice of the individuals. For the interpretation form, investigators make interpretation of the statistical research results, or they interpret the themes or patterns that arise from the research data (Creswell and Creswell, 2018).

<b>Quantitative Methods</b>	<b>Qualitative Methods</b>	<b>Mixed Methods</b>
Predetermined	Emerging methods	Both predetermined + emerging
Instrument-based questions	Open-ended questions	Both open- and closed-questions
Performance data, attitude data, observational data, and census data	Interview data, observation data, document data, and audio-visual data	Multiple forms of data drawing on all possibilities
Statistical analysis	Text and image analysis	Statistical and text analysis
Statistical interpretation	Themes, patterns interpretation	Across databases interpretation

Table 4. 3: Research methods

#### **4.4. Choice of Research Approaches**

As previously stated, clarifying research approach of a study is an important aspect especially for the research readers since it will enable them to know the study's strategies and procedures that span the stages from wide assumptions to specified methods of data collection, analysis and interpretation. Creswell and Clark (2007) point out that it is crucial for a researcher to think carefully about the best methodology for his/her study so as to make certain that the research objectives are achieved and that the findings are validated.

Our research framework elements are the mixed methods approach, the pragmatic worldview, and the mixed methods design. This research selects these components because they suit the nature of our study. The research mainly employs the sequential exploratory strategy as we first start with qualitative phase and we build up our quantitative phase from the outcomes of the first phase. That is to say that, the research method features qualitative data gathered from experts' opinions and focus group discussions. These data enabled the researcher to design comprehension tests to evaluate human ease of reading on selected texts. This study is expected to examining the effectiveness of applying readability metrics as an indicator of policy

comprehensibility and whether the readability has an influence on understanding information security policies (ISPs).

The proposed study is also descriptive and employs statistics in describing the aspects of the various samples that are utilised in the research. It can be seen clearly in the main study stage when the characteristics of a population were described. This research is not only descriptive but also comparative since the quantitative approach results compare with software readability formulae results. In addition, this research does not depend solely on text measurement as many readability researchers do but also complements it with comprehension tests as a confirmation of text measurement. This may afford superior understanding for the use of a readability standard as a basis for predicting likely comprehension, particularly for written text. The following subsections consider the methods as they are utilised in this research.

#### **4.4.1. Qualitative Methods**

The qualitative method is applied to a small-sized sample, yet it makes available the overall comprehension of the issues through a blend of different procedures and methods to get data, for example, intensive interview and focus group discussion (Cohen, Manion and Morrison, 2011).

Following Glassner and Moreno (2013) advice of adopting the qualitative method that can be a most benefit in areas where there is no or little existing knowledge, this research will use qualitative methods for the early stages of our research methodology steps. Qualitative methods have been employed in this dissertation to achieve the objective (identifying and clarifying salient points relevant to ISP documents in order to use the findings for the following stage, quantitative phase) discussed in Chapter One by carrying out and analysing experts' opinions and focus group discussions, as elaborated below.

#### **Expert Insights**

In this study, eight IS policies were selected to examine their readability by programmatic means and human comprehension tests. Without the use of pre-determined inquiries, experts were asked for their insights on those policy ingredients that they considered key. There are various well-known techniques for gathering

expert insight and our strategy was to adopt an easy but effective method (Crosby, 2016). The aim was to benefit from the experts' professional experience in identifying and clarifying salient points relevant to ISP documents.

### **Focus Groups**

A focus group discussion is an effective qualitative method to collect data by selecting a group of people, in which they are requested to express their opinions about predetermined questions (Mazza and Berre, 2007). It is also defined by a number of scholars as a group of participants chosen and assembled in order benefit from their personal experience to discuss and comment on the topic that is the subject of the study (Powell, Single and Lloyd, 1996). Focus group approach normally consists of a small group conversation (usually between five to eight members) organised by a catalyst that has a well understanding of members in addition to attitude applicable to a particular study (Gorman *et al.*, 2005).

Liamputtong (2011) and Nardi (2015) comment that focus group technique is appropriate for studying opinions and attitudes, and it sparks interactions between participants that are likely to enhance discussion and insight which result in providing a researcher with a lot of information. Furthermore, Gorman and Clayton (2005, p. 143) note that 'a variety of perspectives and explanations [to] be obtained from a single data-gathering session'. Following the evaluation research team (2008) advice of employing focus group discussion in combination with another method to clarify and evaluate research findings, focus group interviews were conducted to confirm the expert insight. Therefore, a number of selected participants in an informal meeting were asked to express their opinions on the most salient statements from a number of chosen policies.

It is worth mentioning that, achieving confidence in both steps (experts' insight and focus groups) was considered vital, as the perspectives of the experts and focus group members would later be the basis for determining how well the documents convey these points to less experienced computer literate users. The focus group outcomes provide the information being sought and meet the researcher's expectations. Therefore, the results are productive, usable and crucial for the next stage of our study.

#### **4.4.2. Quantitative Methods**

The quantitative technique can be defined as research including the utilisation of organised questions, with prearranged response choices, that are administered to many participants and the data produced are usually numerical (Neuman, 2013). The quantitative method results are commonly produced from a large sample size and therefore the result can be generalised to a larger population (Scandura and Williams, 2000).

The quantitative method has been used in this research to find an answer to one of our research questions (Can readability, as gauged by software metrics, accurately reflect the comprehensibility of ISPs? If so, how?), discussed in Chapter One by carrying out and analysing questionnaire instrument, as elaborated below.

#### **Survey**

The term ‘survey’ generally means a study that has utilised a representative sample (Pickard, 2013). A survey is a suitable strategy for methodically collecting data from a various scope of individuals with diverse social backgrounds (Schutt, 2011). The survey instrument has many advantages including 1) it is not an expensive method, especially web-based surveys, 2) it can reach larger samples, 3) it can ensure anonymity, 4) it can deal with several topics in one questionnaire, 5) it avoids respondents’ time pressure and interviewer bias, and 6) it is one of the easiest methods to code closed-ended items (De Leeuw, J. Hox and Dillman, 2012; Nardi, 2015).

The survey method is an ideal instrument for our main study phase as it can address a large number of samples in a short time and at a moderate expense. As Briggs et al. (2012) state that the survey tool is one of the most widely utilised research methods, and certainly the most employed quantitative strategy because of the advantages that distinguish it from other methods.

The main study survey was devised and deployed using the Qualtrics survey platform (Qualtrics, 2017). Despite the fact that the Qualtrics software tool was not designed specifically for creating online Cloze tests, Qualtrics was adopted because it could meet our requirements for creating the Cloze test question structure and had many features (see the main study’s data collection section for more details).

## **4.5. Research Design (Exploratory Sequential)**

The exploratory sequential mixed methods design that can be described as the converse of the explanatory sequential mixed methods design begins by collecting qualitative data, thereafter collecting and analysing quantitative data, before reaching the stage of interpreting the results. The intent in using this type of design is either to explore a field (perhaps because variables or hypotheses are not available) or to develop and test a new instrument (Creswell and Creswell, 2018).

One of the main advantages of the sequential exploratory strategy is the straightforward nature of the approach. It is easy to implement due to the fact that the phases fall into clear separate components, besides, the design is easy to describe and the results easy to report. The second notable feature of the exploratory sequential design is its remarkable procedure, which addresses a quantitative phase after achieving qualitative results, and this is ideal for many studies such as our own as our study requires a second quantitative stage emerges based on what is learned from the first qualitative stage.

On the downside, this approach requires a lot of researcher time since it has two separate phases and they have to be addressed in sequence (qualitative and then quantitative). In other words, investigators may spend a lot of time to implement this design properly and then interpret the results. Moreover, it can be challenging for some researchers to adopt the information from the qualitative phase to construct the quantitative phase (called the integration point). Furthermore, there is a risk that researchers might inadvertently develop an inappropriate instrument or measurement that does not take full advantage of the richness of the qualitative findings (Teddlie and Tashakkori, 2010; Ary *et al.*, 2018; Creswell and Creswell, 2018).

### **4.5.1 Why the Exploratory Sequential Approach is Selected**

There are a number of reasons behind our preference for the exploratory sequential mixed methods design over other mixed methods design, see Figure 4.3 that illustrates our exploratory sequential mixed methods design (more details are provided in Chapter Five). Firstly, although the convergent mixed methods approach is a single-phase design combining the quantitative and qualitative data, our study requires a sequential method in order to answer our research questions. The important



assumption in this design is that both qualitative and quantitative data deliver different kinds of information and together they produce a coherent unified result. Therefore, this type of mixed methods design does not fulfil our requirements because of the convergent mixed method design cannot be used to find an answer for our research questions.

Secondly, the explanatory sequential mixed methods design is a two-phase design in which the preliminary quantitative data outcomes are explained further with the qualitative data. If we chose to adopt this design in our study, we would have to start with the survey (quantitative) and then focus group interviews (qualitative). This approach is not viable, since the survey instrument cannot be developed before identifying its questions and the questions are derived from the focus group stage. Consequently, because it starts with quantitative data and follows up with qualitative data, this type of mixed methods design is also inappropriate for our study

Thirdly, the exploratory sequential mixed methods design was chosen because of their many features including A) it is easy to describe, implement and report, B) it can be used when a study requires a second quantitative stage emerges based on what is learned from the first qualitative stage, and C) a researcher can create a new instrument as one of the possible products of the study process.

Fourthly and most importantly is the exploratory sequential mixed method approach is the most suitable design for our research due to the nature of our study. It enables us to collect qualitative data, which is used later for defining the Cloze test questions, and then these questions are used in our quantitative method (survey) stage. Furthermore, the outcomes of the exploratory sequential approach will be compared against a number of readability formulae to identify the degree of similarity and difference between the results of human and software metrics. This procedure cannot be achieved with other designs, (e.g. convergent mixed methods or the explanatory sequential mixed methods designs) because we need to establish the results from the initial stage before moving to the final comparison. The overview of our study framework can be seen in Figure 4.4.

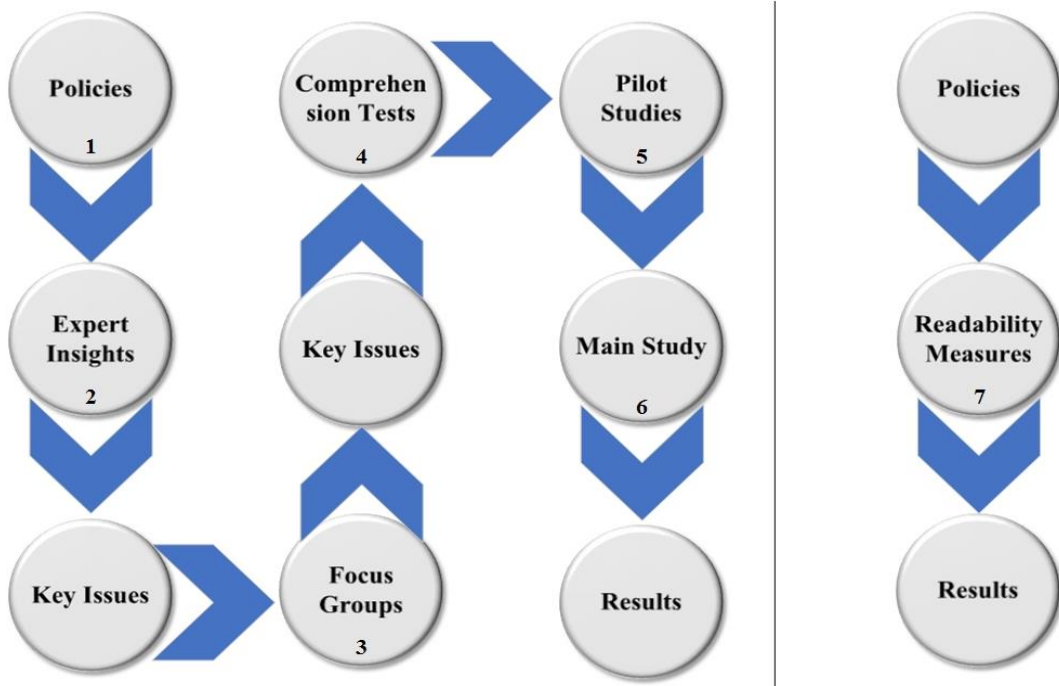


Figure 4. 3: The exploratory sequential design for the thesis

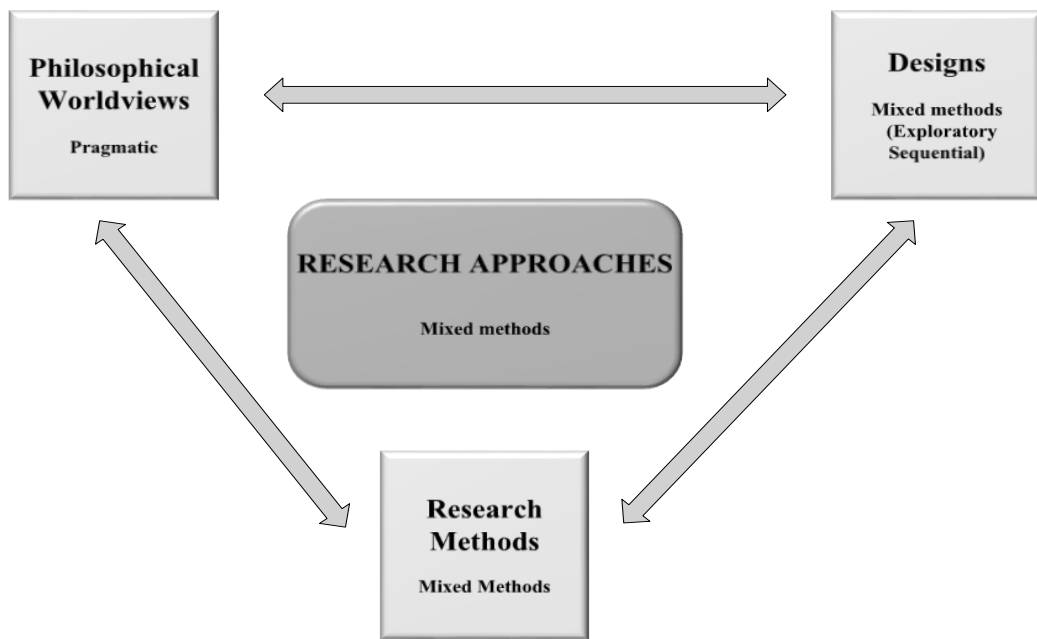


Figure 4. 4: Our research framework

## **4.6. Ethical Assurance**

Obtaining essential permissions to study participants is crucial for a researcher prior to a study. Pickard (2013) and Bell (2014) insist that every single participant have the privilege to know they are being studied and why. Therefore, all ethical instructions and guidelines were seriously considered. A summary of the research including overview, aims and objectives were written in ethics application form. Besides, explained of how recruiting participants and how demonstrating consent. Participants were assured that all processing of data would adhere to the Data Protection Act 1998.

The participant's contribution and the conduct of the research was clearly explained to potential participants. All participants were informed about the confidentiality of the data collected, and their names would remain anonymous. Participants had the option of partaking or decline partaking in the study, and they had an option of accepting or rejecting to be recorded. The informed consent form was written by following the University's Code of Practice on investigations on Human Beings as well as following a set of elements that introduced by Sarantakos (2012). The email was used as a means of securing consent from participants.

Storing data in a secure environment is a top priority of the researcher and the data are only be accessed by an authorised person. After the empirical data had been compiled, the qualtrics.com account was deactivated. The data would be destroyed comprehensively and securely when the researcher finishes his study. Creswell (2013) explains that surveys and data have to be disposed of them after a retention period of five years.

The study was compiled with a full consideration of all standards for ethical research. To inspect and solve any ethical matters concerning human rights violations and to protect the participants' privacy and anonymity, two ethical application documents were submitted to the ethics committee of the Computer and Information Sciences (CIS) department at University of Strathclyde. Creswell (2013) states that a researcher requires displaying his research plans to be reviewed by an Institutional Review Board within the researcher's organisation.

## **4.7. Summary**

The exploratory sequential mixed methods design was utilised to answer the study question of this research. The following data collection methods were adopted: conducting experts' insights, holding focus groups, and distributing questionnaires. The next chapter will discuss in detail our sequential exploratory strategy, as well as, a number of pilot studies will be presented.

# CHAPTER FIVE: EXPERIMENTAL DESIGN

## 5.1. Introduction

This chapter provides a comprehensive description and justification of each phase used in the exploratory sequential mixed methods design. This includes purposes, procedures, participants, and the results of each phase that employed in the sequential exploratory strategy. This chapter also focuses on the selection and justification of the sample size undertaken in this research as well as presenting the data collection of both approaches (qualitative and quantitative).

## 5.2. The Research Methodology Steps

This research was conducted in seven phases to achieve the study objectives. Each phase was developed in light of the outcomes from the previous phase. The figure below illustrates the steps of these phases.

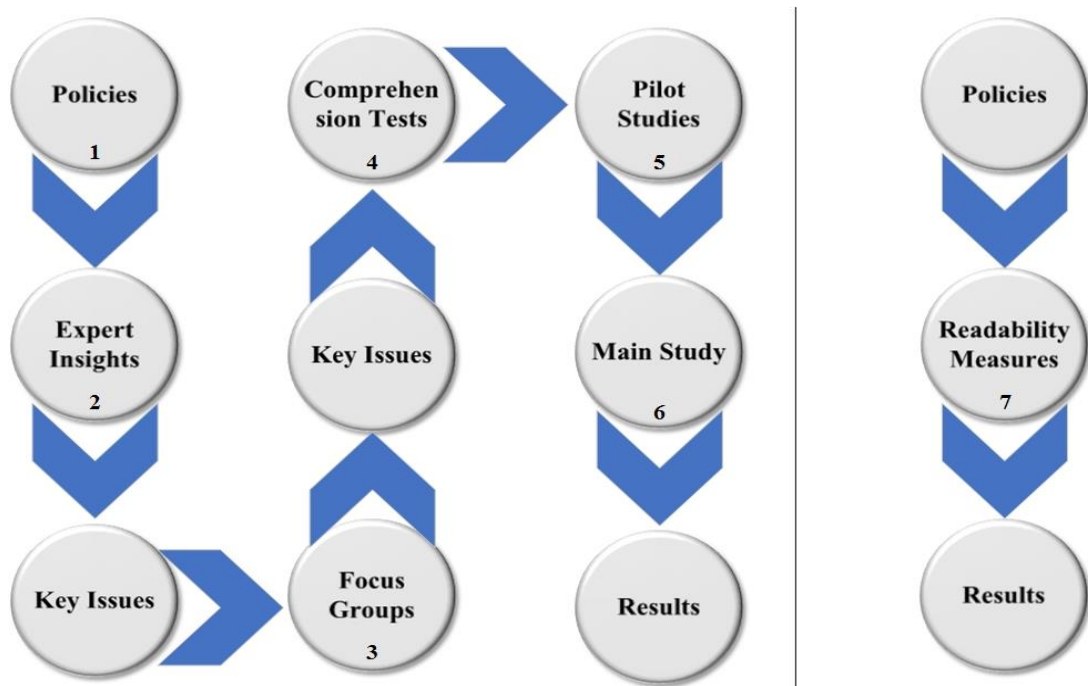


Figure 5. 1: The research methodology steps

### **5.2.1 Phase1**

In the first phase, eight information security policies (ISPs) were selected (from a larger set of thirty-five policies) to examine their readability by programmatic means and human comprehension tests. The chosen policies are a mix of public and private sectors (academia and industry). Five of the ISPs are universities, and the others are telecom organisations, in order to add a further comparative dimension and determine whether these sectors are similar.

These policies were not chosen randomly but were carefully selected based on matching a number of factors. Policies have to be from countries where English is the mother tongue. In addition, the policies should be from a variety of geographical locations, accessible online and with similar word count (no more than 10% difference). In addition, University policies were chosen from top universities of 2015 - according to Quacquarelli Symonds (QS).

### **5.2.2 Phase2**

In the second phase of this work, without the use of pre-determined inquiries, experts were asked for their insights on those policy ingredients that they considered key. There are various well-known techniques for gathering expert insight and our strategy was to adopt an easy but effective method (Crosby, 2016). In this context, the expert is someone who has worked with policies for at least 15 years. Seventeen responses were received from this experts group. The aim was to benefit from the experts' professional experience in identifying and clarifying salient points relevant to ISP documents.

Achieving confidence in this step was considered vital, as the perspectives of the experts would later be the basis for determining how well the documents convey these points to less experienced computer literate users.

### **5.2.3 Phase3**

In Phase3, we conducted focus group interviews to confirm the expert insight. A number of selected participants; in an informal meeting, were asked to express their opinions on the most salient statements from a number of chosen policies. Focus group

discussion has a number of valuable features including 1) it can enable comprehensive discussions and involve a small number of participants, 2) it concentrates on a precise area of interest and enables people to discuss an issue in depth, 3) it sparks interactions between participants that are likely to enhance discussion and insight (Liamputtong, 2011). In addition, focus groups can be used in combination with another method to clarify and evaluate research findings (Evaluation Research Team, 2008). For these reasons, focus groups were used to validate the views expressed by the computer experts in the previous stage. Thereby, we have adopted two qualitative method approaches to determine the validity of the result before proceeding to the following step.

The perfect size of focus group discussion is a contentious subject. The size of the focus group was determined as between five to eight and the discussion time between 60 to 90 minutes. In order to make certain that participants have the opportunity to share their views without getting bored with the process. This decision follows Krueger and Casey (2014). Eleven people (professional computer users, who have used computers for more than six years), who expressed an interest in the study, were invited to participate in a 90 minute discussion. This approach similar to the focus group meetings described by Bray, Johns, & Kilburn (2011), and Yunos, Hamid, & Ahmad (2016).

#### **5.2.4 Phase4**

Phase4 focused on developing comprehension tests (Cloze tests). Following Bormuth (1967), Guillemette (1989), Kobayashi (2012), Rankin & Culhane (1969), Richards & Schmidt (2015), Taylor (1953), these Cloze tests are key to determining the comprehensibility of policy components and underpin both the pilot and main studies.

As noted, several readability formulae have been explored in the literature as a basis for gauging the readability of written material, but invariably depend upon syntactic variables (Klare, 1985; DuBay, 2004; Anagnostou and Weir, 2006; Campbell and Weir, 2006; Crossley *et al.*, 2017). DuBay (2004) advised in his inclusive review on readability researches, that future studies about readability should consider the limitations of readability metrics by utilising other methods together with the

readability formulae. Thereby, comprehension test will be adopted in conjunction with readability formulae to measure the difficulty of the selected text materials.

#### **5.2.5 Phase5**

In Phase5, as a pilot study, a number of Cloze tests were evaluated prior to the full-scale study. As noted by Connelly (2008), the main aim of performing a pilot study is to field test logistical characteristics of the upcoming study and to include these aspects of the survey design. Besides, The Scheaffer et al. (2011) point out that pilot study outcomes mostly propose some changes that necessity be made before the main study is launched.

This step is essential and beneficial in establishing the groundwork in a research study and can save a considerable effort, time, and money by identifying potential issues and inadequacies in the examination instruments before embarking on the primary study (Abu *et al.*, 2006; Pickard, 2013).

#### **5.2.6 Phase6**

The sixth phase is the main component of the research. Self-administered questionnaires were created by the popular Qualtrics platform and widely distributed. The survey method is an ideal instrument for the main study phase as it can address a large number of samples in a short time, it is also cheaper than other methods such as telephone interviews or focus groups interviews (Nardi, 2015). The term ‘survey’ generally means a study that has utilised a representative sample (Pickard, 2013). As a side note, for our study, the term ‘survey’ and ‘questionnaire’ mean the same thing and they will be used interchangeably.

According to Bryman (2015) deciding on the suitable size of a sample is not straightforward and there is no definitive answer. However, A review of the literature concerning valid and reliable sample size found that scientists suggest the minimum number of sample is 200 respondents and the most suitable size is 384 survey questionnaires (more details provided in the sample size section) (Krejcie and Morgan, 1970; Harris and Schaubroeck, 1990). A total of 600 people participated in the study. Responses from two hundred and five respondents were later withdrawn on validity



issues. The remaining data of three hundred and ninety-six subjects were entered into the Statistical Package for Social Sciences (SPSS) program for analysis.

### **5.2.7 Phase7**

The final phase compared the comprehension test outcomes against readability metrics to identify the degree of similarity and difference between the results of human and software metrics.

## **5.3. Information Security Policies Documents**

Eight IS policies were selected from a considered set of thirty-five policies to examine their readability by means of software readability metrics and human comprehension tests. These policies are a mix of public and private sectors (academia and industry). Five of ISPs are universities, and the others are telecom organisations in order to add a further comparative dimension and determine whether these sectors are similar. The universities are from several countries (Australia, United Kingdom, and United States) as well as companies from different countries (Australia, Canada and United States), but I assume they cover similar contents. The reasons behind choosing policies from various countries to accommodate the possibility that the local forms of English would affect the ease of reading.

These policies were not chosen randomly but were carefully selected based on matching a number of factors. Policies have to be from countries where English is the mother tongue. In addition, the policies should be from a variety of geographical locations, accessible online and with similar word count (no more than 10% difference). In addition, university policies were chosen from top universities of 2015 - according to QS website. Searching for these policies' documents was done manually utilising an Internet search engine.

### **5.3.1 Core Process**

This core process explains the steps adopted in selecting and preparing the ISPs documents.

1. Search for ISPs: this step is for university policy. It starts by reviewing (from the top to down) the presented universities by QS website. Then, discard universities that their first language is not English to avoid occurring errors in translation. After that, picking universities taking into consideration the balance in the number of universities from different regions.
2. Examine the selected policies: checking if they have online IS policies, if not it would be dismissed. For private sector policy, this stage was the most difficult step in the process because a very few English ISPs were published online.
3. Compile and modify the policies: for each policy, information relating to ISP would be collected; so, all information that was on many web pages would be gathered in one document. Then, removing unnecessary information such as links, contact information, dates..., etc.
4. Approve or reject policies: the number of words in policy documents was calculated and documents with similar word count were selected, since a different word count may skew the readability (maximum of 10% difference was accepted).
5. The last touch: reviewed policies and selected five universities' policies representing the public sector and three of telecom organisations as the private sector, with taking account of geographical diversity.

Therefore, this study contains five ISPs for the public sector, which are Cambridge University (5434 words), Imperial College (5247 words), Melbourne University (5609 words), Princeton University (5501 words), and Stanford University (5672 words). In addition, three ISPs for the private sector, these are Telstra Mobile Company (5507 words), TELUS Mobility Company (5842 words) and T-Mobile Company (5374 words). These policies will be used as ingredients in our experiments, and they will be placed in various readability measurement tools.

## **5.4. Prior to Experts' Insights Experiment**

### **5.4.1 Preparing Draft Instrument**

Pre-testing is an essential and important requirement that should be seriously considered before starting to distribute a survey to such a large group of volunteers. De Vaus (2013, p. 48) asserts that it is judicious to evaluate the reliability and validity of indicators earlier than conducting the actual questionnaire with the entire sample.

One of the purposes of doing pre-test is to explore certain issues before undertaking a large-scale study, such as testing feasibility in practice, evaluate or purify the questions and enhancing the methodological quality of the experiment's parts. Hence, it helps the researcher to discover and solve difficulties prior to the primary study.

### **5.4.2 Procedure**

An introductory email (see Screenshot 5.1, Appendix B), that includes brief information about the researcher and about the nature of research, volunteers' task and some guidelines and instructions, invited participants (computer professionals) to contribute to the pre-test.

Another email (see Screenshot 5.2, Appendix B) was sent to those who replied positively to the first email. The task for each pre-test participant is to read two policy documents and give feedback on what they consider being the most salient points in the content. The policy documents were attached in Microsoft Word format to be easier for the participant to rank, highlight and add comments. The estimation of completing one report would require about 15 minutes.

### **5.4.3 Participants**

As this research would involve human participants, ethical approval was required, as listed in Appendix B. The pre-test started once the ethical approval has been obtained.

In this pre-test, twenty PhD students from the Department of Computer and Information Sciences, University of Strathclyde, Glasgow, were invited to take a part. Ten out of twenty individuals were willing to participate; therefore, a replied email was sent to them with an attached assignment.

#### **5.4.4 Results**

The participants said they spent more than the expected time to finish a single task. Overall, respondents were able to understand most of the guidelines and instructions to answer the questions; only a few of them experienced difficulties in that. After the respondents' comments, the self-administered survey was enhanced by presenting the instructions in an easy video clip to enable participants to answer the question correctly (Alkhurayyif, 2015).

#### **5.4.5 Summary**

As pre-tests show a fundamental stage of the study process, a pre-test on investigating the effectiveness of adopting readability metrics on eight selected policies was conducted prior to doing the actual study. Thus, obtaining valuable benefits from the feedback of the pre-test's sample and enables last-minute corrections and adjustments. Specifically, facilitating the task's guidelines and instructions, increasing the expected time to finish the work, expanding the amount number of participants and limiting the assignment of each member to one policy instead of two.

### **5.5. Experts' Insights Experiment**

A number of professional users, who use the computer more than six years, validated and tested the questions and made certain that the questions could be understood, before taking the step of asking experts about their opinion. In this context, the expert is someone who has worked with policies for at least 15 years. There are various well-known techniques for gathering expert insight and our strategy was to adopt an easy but effective method (Crosby, 2016).

The aim of using this method is that obtaining valuable benefits from experts' professional experience or education in identifying and clarifying salient points relevant to ISP documents. This step of the research is critical, as it may cause the study process to be extended if we have not received enough respondents.

### **5.5.1 Procedures**

The same survey used in the pilot study was employed with the experts consulted. For the aims of this piece of work, ‘experts’ in this sense may involve an individual who has experience in policies more than 15 years of industrial or education sector.

An introductory email was sent during October 2015 to non-random selection of participants. Two weeks later, a reminder email was sent to people who have not replied yet. The email contains information about the researcher, research, participant’s contribution task and the study purposes. Hence, illustrating precisely the research objectives and the expectation from experts, which are important to achieve satisfactory results. In addition, it was stated in the email the human right protection and the benefits of participating. They were informed that participation in this study is voluntary, and they were under no obligation to take part in this study. Furthermore, they could withdraw from part or all of the study at any time without any consequences. Besides, any information recorded during the study would remain confidential as well as participants’ names will not be used and the survey outcomes would be reported in the aggregate without reference to any identifiable individual. Last of all, members required signifying their consent by proceeding with an email before getting an assignment.

A second email was sent to those who replied positively to the first email as shown in appendix B. It included guidelines and instructions on how the participants could understand the task and completed it fruitfully. In addition, the email included contact information and a one policy document to be analysed. After receiving the responses, participants were thanked for their appreciated time and exertion in joining in the study. The steps to design and probe the experts’ insights experiment are shown on Figure 5.2.



Figure 5. 2: Experts insight experiment process

### 5.5.2 Participants

Before conducting this study, the consent was taken from the ethics committee of the computer and information sciences (CIS) department at Strathclyde University, and additionally from the members themselves. The researcher aimed to distribute the questionnaire to more than forty experts through emails in order to get replies from half of that number. Inviting a large number of experts to contribute to the study was not an easy task. To start with, the head of Strathclyde University infrastructure proposed some potential participants along with the researcher’s supervisor. Others were contacted directly by the investigator.

In principle, there are no particular requirements for the participants, aside from them having some expert knowledge of the area under discussion. Forty experts, who have a different position in IT departments of various organisations, received email requests for survey participation (see Screenshot 5.1 & 5.2, Appendix B). A total of seventeen separate email correspondences were received, nine from educational

background, seven from telecom organisations and one from military sector. It is worth mentioning that every participant joined in the study on a voluntary premise.

### **5.5.3 Conclusion**

To conclude, only seventeen feedbacks were received from the members, which meant that each ISP document received two expert opinions with the exception of the Melbourne University policy document, which had three comments. The rate was below 50%, which could be attributed to one or more of the following reasons. Some experts may lack interest in taking part in the study or busy at that time. The sensitive nature and/or complexity of the policies might have appeared intrusive and restricted expert response. The elucidation that the participants' responses were classified and data analyses were anonymous might not have overcome individual apprehension of associating email addresses with study responses. The individual effort required to read a policy document, select important points and rank the selected points from most to least important might have appeared to be too over-whelming or complicated for participation. Last of all, experts doubt over the time required to read fourteen pages and select the salient points may have allowed other priorities to take precedence over the study.

To minimise the effects of low rate of responses, a decision was taken to employ an additional stage in which the focus group validated feedback from experts. Achieving confidence in this step was considered vital, as the perspectives of the experts would later be the basis for determining how well the documents convey these points to less experienced computer literate users.

## **5.6. Focus Group Methodology**

A focus group is an effective qualitative method to collect data by selecting a group of people, in which they are requested to express their opinions about predetermined questions (Mazza and Berre, 2007). It is also defined by a number of scholars as a group of participants chosen and assembled in order benefit from their personal experience to discuss and comment on the topic that is the subject of the study (Powell, Single and Lloyd, 1996).

A number of selected participants; in an informal meeting, were asked to express their opinions on the most salient statements from a number of chosen policies.

Focus group discussion has a number of valuable features including 1) it can enable comprehensive discussions and involve a small number of participants, 2) it concentrates on a precise area of interest and enables people to discuss an issue in depth, 3) it is appropriate for studying opinions and attitudes, and 4) it sparks interactions between participants that are likely to enhance discussion and insight (Liamputtong, 2011; Nardi, 2015). In addition, focus groups can be used in combination with another method to clarify and evaluate research findings (Evaluation Research Team, 2008). However, this technique can be difficult to new or inexperienced researchers since the researcher roles of managing focus group discussion as well as keeping the session on track. The focus group organiser acts as a mediator between the question and the group and between an individual participant of the focus group (Creswell and Creswell, 2018). For the introduced features of the focus group, this approach was used to validate the views expressed by the computer experts in the previous stage. Thereby, we have adopted a two qualitative method approach to determine the validity of the result before proceeding to the following step.

The perfect size of focus group discussion is also a contentious subject. The size of the focus group was determined as between five to eight and the discussion time between 60 to 90 minutes. In order to make certain that participants have the opportunity to share their views without and getting bored with the process. This decision follows Krueger and Casey (2014).

### **5.6.1 Focus Discussion Group Report**

The meeting was held on January 28, 2016. A select number of participants were carefully recruited, representing the interest in the case study and invited to the focus group discussion. Eleven people were invited to participate in a meeting. The single parallel session ranged from six participants in the first group and five participants in the second group and lasted for 90 minutes. This approach is similar to the focus group meetings described elsewhere (Bray, Johns and Kilburn, 2011; Yunos, Hamid and Ahmad, 2016).



Each group discussed, selected and ranked the most important statements of their task through a dynamic exchange of ideas among all the participants. The moderator posed eight questions, such as the following:

**Rank 10 of the following statements of Policy A to indicate their importance to you (1 is most important, and 10 is least important).**

Focus group questions were developed with direction and supervisor feedback and were designed to cover and verify the result of the previous step. The focus group was assembled to characterise the core of each policy, which would contribute to the next stage of our study.

The focus group was conducted as part of an experiment to determine whether the readability has an influence on understanding ISPs. Participants provided information in two methods: written responses and group discussion.

The rest of this section will clarify the preparation of focus group meeting and present the focus group methodology. Subsequently, the results will be presented before highlighting the implications in the discussion.

### **5.6.2 Procedure Preparation**

The arrangement of focus group meeting starts immediately after completing the expert insight stage. The purpose of adopting group discussion is clear (to validate what computer experts suggested in the previous stage), and the project work plan was developed. Hard copies of the questions; which it is ranking and selecting ten statements of eight policies; were printed out to be distributed to focus group members later.

The preparation of focus group started by determining the most suitable people to participate in this step. The target people are professional computer users, who use the computer more than six years such as CIS staff, researchers and postgraduate students (Masters and PhD). Then deciding the holding place, time and proposing two different dates. The researcher reserved a quiet room at the computer and information sciences department to be familiar location and easily accessible to interested members. Besides, two laptops were booked so that the potential participants can easily access the full version of the policies if required.

Lunchtime was chosen to be the right moment for the meeting as staff and researchers usually have their lunch break, students usually do not have lectures and a moderator could serve some snack foods and beverages for attendees. As stated before, two different dates were offered to interested participants in order to select the most convenient time for them.

### **5.6.3 Methodology**

Advanced planning was required to ensure that potential participants can be brought together at appropriate times. An initial email and later a reminder email were sent to the mailing list of CIS staff, researchers and postgraduate students of the Strathclyde computer and information sciences department (see Screenshot 5.3, Appendix B). The researcher intended to send the first email on Sunday because it has the highest open rate of opening messages according to many recent surveys (Vanderkam, 2012; VerticalResponse, 2017).

Sunday has now become the best day to send an email because many people check their email on Sunday, fewer messages come in, and if it was not read on Sunday, it would likely be the first thing read on Monday morning. A reminder email was sent on Tuesday, also considered by some as the best day to send an email (Neely, 2013; Ellering, 2016), contained: a mini introduction of the researcher and topic, place, time, date, expected duration and highlighting participation incentives (prizes and lunch).

Eventually, eleven people were keen, eight PhD students and three Masters level students. All eleven were full-time students. The participants were from different nationalities (Bulgaria, Greek, Saudi Arabia, Libya, Oman, and United Kingdom) which might be a positive impact on the group discussion, as diversity will generally lead to more productiveness, creativeness, and innovativeness (Fine and Handelsman, 2010). The reason for using the smaller group of people was to provide everybody with the chance to express their perspectives and opinions.

Each person in the focus groups was reminded the day before the discussion group. Afterwards, they were divided deliberately into two groups based on their years of experience. Six participants in the first set (four PhD students and two Master students) and five members of the second set (four PhD students and one Master student). Each

group was given four random policies (out of eight policies), but at least one telecom company policy, and they were asked to answer the questions without sequential order. Besides, notebooks paper were provided so the group members can simply write their discussion ideas throughout the group brainstorming session.

The main output of the focus group discussion was highlighting ten most important points in each of the given policies. A moderator simplified the focus group meeting by delivering guidance to the group members and encouraging participants to speak freely and spontaneously about a given subject. He took into consideration advice from Bray et al. (2011) by starting the discussion topic without presenting any perspectives throughout the talk-in-process session. Furthermore, the facilitator strived to create a free-streaming discussion during focus group meeting with least moderator intercession. Furthermore, Bryman (2015) point outs that there is a level of facilitator involvement. A facilitator should allow latitude to participants, consequently, the group discussion can range fairly extensively.

The facilitator took into consideration that the first few moments in the focus group meeting are crucial. Therefore, he started the meeting with a brief PowerPoint presentation that introduced the researcher's information and interest, study overview, participants' task and some guidelines and instructions. Then, allowed a few minutes for questions or concerns. After that, distributing worksheets to each member of the groups. Each group had four policies to answer them. Participants were organised their time efficiently by following a number of steps. First, located a few minutes to read each statement of policy. Then, focus group members could ask their group if they do not understand some statements. In the absence of clarity the group's members needed to look at the full policy to obtain a better understanding, which happened in one occasion. After that, the participant selected and ranked only ten statements based on their opinions. Later on, the leader of the group reads all the policy statements; point by point, and simultaneously put the number of participants' votes per statement. Obviously, some of the points did not have any vote, as there were more than ten statements to be chosen from, (it was between 15 to 20 statements). Thus, the ten statements would be that holding the majority number of votes. After finishing the discussion, participants were thanked for their valuable time and effort in taking part in the focus group meeting.

The moderator wished to record focus group discussion on an audio tape recording to obtain complete discussion, idea and accurate information but unfortunately, some participants declined to have the meeting recorded. Due to that, the moderator took notes throughout the discussion along with other roles such as guiding the debate, listening to participants and timekeeping, which were after members accepted to document their responses in writing, i.e., the discussion ideas that were noted on the notebooks paper were collected at the end of the meeting. See Figure 5.3 that illustrates the focus group process.



Figure 5. 3: Focus group process

#### **5.6.4 Results**

The researcher obtained lots of information, opinions and suggestions as participants were interested in the topic and they were motivated to express their views freely. It is worth saying that focus group members demonstrated their choices with supporting evidence.

After observing both groups, the first group's members did not see any of the full policies of their tasks. As they believed that, each policy statement has adequate information to express the meaning. While, the second group's participants had

checked some ISP statements of Cambridge University. In addition, members of the first team finished their discussion a bit earlier than the second group members; however, the second group participants were more enthusiastic and interaction than the first group.

The majority of the first group members felt that the ISP of Melbourne University was complicated, while the ISP of TELUS Mobility Company was easy to understand. Whereas most of the second-panel participants considered the ISP of Cambridge University as difficult to comprehend, while the ISP of Stanford University was straightforward to read and understand.

Many participants acknowledged that some policies emphasizes more on privacy, security and business aspects. Moreover, some policies cover lots of issues even about emergency plans in detail.

Due to not tape-recording the focus group interview, a number of questions were asked to each group at the end of the discussion in order to ensure that the moderator covered all aspects from the session. The worksheets used in the discussions were transcribed and analysed along with the facilitator notes. Then, the focus group results were compiled in one document for ease of use when required.

The focus group outcomes provided the information being sought and met the researcher's expectations. Therefore, the results were productive, usable and crucial for the next stage of the experiment.

### **5.6.5 Conclusion**

In this study, I have introduced the focus group methodology. Then, the definition, purpose of use and the main features of a focus group meeting have been discussed. This was followed by explaining the procedure preparation, the experimentation implementation and clarifying the outcomes.

Following on from this constructive dialogue, the focus group interviews revealed several key insights associated with the discussed topic. The results include 1) determining the salient points for each policy, 2) insight on the extent of comprehension and ease of reading for a number of the procedures contained in the

policies, 3) indication of which policy of the set was most complicated, and which one was the easiest policy to understand 4) highlighting the variation between policies with respect to the aspects they covered.

## **5.7. Pilot Study of Cloze Tests**

A pre-test is a rehearsal of the main study, which has to be done first in order to expose any difficulties and address them (Higgins and Compeau, 1995; Milne, Orbell and Sheeran, 2002; Alreck and Settle, 2004; Sonderegger and Sauer, 2010). There is a debate among scholars about the number of participants recommended for the piloting stage. Albert et al. (2010) believe that a pilot study sample size should be 10% of the total potential contributors to the actual study. In contrast, Johanson and Brooks (2010) argue that many complexities are surrounding the issue of determining an ideal number of participants for a primary study because many factors influence it. Nonetheless, Hill (1998) and Isaac and Michael (1995) declare that 10 to 30 respondents should be sufficient for piloting a survey research; whereas, Julious (2005) and Van Belle (2011) recommend 12 participants as a minimum. Therefore, the researcher believes that 12 to 30 observations should be adequate for the current stage because of the main study sample size is about 380 participants.

### **5.7.1 Aims**

Prior to the primary study, some preliminary Cloze tests were performed to ensure the practicality of the approach and to enhance the relevance, clarity, and content of the tests. Furthermore, the pre-test would detect possible drawbacks in the proposed approach. Connelly (2008) believes that the main aim of performing a pilot study is to field test logistical characteristics of the upcoming study and to include these aspects of the survey design.

The choice was made to develop an initial Cloze test with a limited number of questions to secure more participants (Mendes, Mosley and Counsell, 2003). A number of variables have been considered, including text length, clarity of instructions, ease of understanding questions, deletion ratio, the nature of words to be restored, blank length, the number of occurrences of restored words, and all words start with a lower-case letter.

As this was a pilot study, the potential participants were encouraged to comment on the pre-test questions. It was hoped that the respondents would propose areas where additional control would be necessary. In addition, the purpose of the pilot study is not entirely to present the results but is to refine the study instrument as much as possible.

### **5.7.2 Participants**

A total of twelve United Kingdom university students, seven males, and five females were recruited in the pilot study. All participants were informed that individual's identities would not be revealed. Their highest level of education was a master's degree. Eight of the twelve respondents whose ages were between 26-34 years old and four people, their ages were between 35-54 years old.

Two study forms were constructed for the pilot test. Each participant was randomly assigned one of the two tasks. These comprised four Cloze tests and four multiple-choice questions (see 5.7.4 Procedures for procedure detail). Twelve of the fifteen people, who were contacted by email, agreed to take part in the initial study. Receiving twelve responses with beneficial feedbacks was adequate for the current step and prepared to proceed for the following stage.

### **5.7.3 Materials**

Reading passages were selected from the set of ISPs from native English-speaking countries such as Australia, Canada, United Kingdom and United States. Policies were selected from a variety of countries to accommodate the possibility that local forms of English might affect the ease of reading.

Eight policies were selected from a considered set of thirty-five policies. All eight policies were given to a number of expert users, who have worked with policies for over 15 years. These specialists were asked to read one of the eight ISP documents and give feedback on what they considered the most salient points in the content. After that, the role of the focus group was to choose 10 statements from the points proposed by the experts.

As noted, this research adopted rational deletion (rational Cloze) rather than fixed ratio deletion since the main concern was to focus on 'key' components in the meaning of the sentences. Choosing fixed ratio 'n<sup>th</sup>' approach could select function words or

other less significant aspects. For this reason, the decision was taken to avoid deleting proper nouns and numbers. The number of items for the Cloze test was set at 20, two items per ISP's statement. Although more than 20 items were possible with some policies, a number of blanks were made consistent across ISPs in order to make it easier for prospective participants to answer and to be easier for the researcher to analysis lately. In total, there were 160 items to be assessed (20 blanks by 8 ISPs). At the end of each Cloze test, there was a box of missing words plus four additional words, to make it harder to guess the omitted words. Participants had a choice of using the provided options or not. After completion by the test subjects, all of the filled gaps were analysed and examined.

The study aimed to test comprehension of the primary ideas, as opposed to details, of the statements. Therefore, the acceptable word scoring method was adopted. This method was introduced earlier in the readability chapter (see Chapter Three, section 3.3.2).

#### **5.7.4 Procedures**

In the Cloze test pilot study, all eight of the chosen ISPs were used. The policies were anonymised to prevent any biasing influences on the participant's responses. Because it was impractical for each participant to take all eight Cloze tests, the eight ISPs were divided into two forms. Each form contained four out of the eight selected ISPs, and these four policies were randomly selected by taking into consideration that at least one telecom company policy included in a form. Each form was divided into three sections: starting first by instruction and general information about the study, then four demographic questions (multiple-choice questions with single answers), and finally four Cloze tests.

The Cloze test pilot study started by disseminating one of the two versions to participants in hard copies. The pilot study was used to test the construct validity of item responses, and to count the time spent on the entire survey. Roughly, 25 minutes were estimated for completing the self-administered survey. However, the participants were urged to record their time to know the actual time spent for the survey. Once all of the participants had finished, their responses were marked and analysed by the



survey designer. In addition, the acceptability of the provided answers was checked with assistance from an educated English native speaker.

### **5.7.5 The Pilot Study Results Implication for the Main Study**

The original draft of the survey included ten questions, six about demographics and four Cloze tests. The refinement procedures of face validity (e.g. no obvious typographical or syntactical errors) and pretesting reduced the questionnaire to eight questions for the first pilot study, four for each of demographic question and Cloze test question.

Education level was not deemed an appropriate question for the pilot study's target sample of postgraduate students. The refinement process generated an "other" response with text entry option included with some of the demographic questions. There are some lessons have been learned during and after performing the pilot study. The result of the pilot test and suggestions for the primary study are summarised as follows:

1. *Changing the strategy of survey distribution- Participants perspective.*

Initially, the questionnaires would be spread on web-based surveys and traditional paper surveys. After listening to the pilot study takers' feedback, online surveys will be sufficient as participants in average spent more than 45 minutes to complete the questionnaire. A questionnaire of this length logically will not receive a significant amount of responses, which is a critical factor. Another reason for not distribute the survey on paper (hard copy) is the design limitation. Paper questionnaires cannot provide drag and drop feature that makes the process easy to use and appealing to the prospective participants.

2. *Changing the strategy of survey distribution- Examiner perspective*

Web-based surveys will save examiner time and effort. The sample size for this study is more than 380 questionnaires, which require a large amount of time to be analysed so by using online survey method, the questionnaire's responses will be analysed and stored automatically in database or survey software system. Furthermore, the examiner will no

longer require correcting participants' responses; he only needs tiny time to collect data. Besides, using web-based questionnaires can help eliminate vague, unclear and spelling mistake answers, which lead to data that are more intelligible. Moreover, sophisticated online survey software provides real-time analysis tools so the surveys progress can be tracked.

3. *Maximise the response rate of the survey*

The study survey will be online to reach more participants. Many web-based survey tools enable the survey creator an option of marketing the survey on several online platforms, such as emails, social networks, and a personal website or blogs. Besides, online surveys are more convenience for responses as it allows potential participants to access the survey where and when the respondent prefers, authorising a respondent to save their questionnaire and return later. This flexibility is lacking in paper-based surveys.

4. *Reduce the burden on prospective participants*

In response to feedback, a number of questions and options given have been restated, omitted or modified either linguistically or structurally. The type of Cloze test questions were changed from open blank with the option below to be typed in to drag and drop type, to avoid grammatical and spelling mistakes and to reduce the completion time for the survey. In addition, dispensing the idea of adding four extra options to the options given to the participants. Due to the facts that the aim of using it has dwindled after adopting randomisation on the options provided, and it consumes lots of respondents' time and effort. Moreover, the instruction section has been modified and shortened to be clearer and more precise. Furthermore, reducing the number of items required to be filled per policy from twenty to ten items. Dropping ten items for each policy imply that forty items will be deleted for the entire survey, which raises the chance of receiving more responses that are complete.

#### 5. *Inserting further questions and features for the survey*

A few additional questions have been added to reflect study aims and objectives. In addition, adopting advanced logic features that only web-based survey software contains. For example, adding features to customise the experience for the participants, including randomisation and piping answers into subsequent questions.

#### 6. *More incentives, more responses*

In the beginning, the researcher suggested a donation to the United Kingdom cancer research for the first fifty completed questionnaires to encourage potential participants. A minor adjustment was made in response to the comments from the pilot study group, which instead of the donation, two prizes of gift cards will be randomly drawing. The prize-winners would be notified by email.

### **5.7.6 Conclusion**

The purposes of performing this pilot stage were to test the survey in assessing its feasibility and the effectiveness of the incentives to attract more participants, gathering preliminary data, verifying the clarity of survey instructions, identifying potential problems in survey design, structure and logistic, measuring the consumption time for filling the questionnaire, ascertaining the survey questions reflect the desired end of adopting the questionnaire method, and also the consistency and the rational sequence of the survey items.

In brief, the pilot study lasted for two weeks. Twelve pilot surveys and forms were completed and their details inserted into an Excel sheet; they were analysed, and it can be summarised that:

- All the variables likely to affect the study aims and objectives were partially covered with the percentage of (92%), whereas 8% (Moderate degree) were asking two additional demographic questions in the questionnaire to gain more information about potential participants.
- Twelve out of fifteen distributed questionnaires have received in two weeks' time. The participation represents a fairly high response rate, which is 80%.

- The instructions to proceed in filling the survey were very satisfied with a high percentage (100%) as revealed in all the pilot responses.
- There was neither comment nor complaint about the sequence of the survey items.

Based on the results above, the researcher validated the instruments of the survey for data collection at this phase of the research study.

## **5.8. Second Pilot Study**

### **5.8.1 Aims**

The pilot study was intended to explore various issues such as ensuring the practicality of the research and to enhance the relevance, clarity, and content of the tests. Furthermore, the pre-test would detect possible drawbacks in the proposed approach. The purpose also of the pilot study was to detect some sign of the connection between the variables in the survey questions and reading comprehension. Moreover, Creswell and Creswell (2018) highlighting the importance of pilot testing as it helps in enhancing questions, format, instructions, as well as recognising possible concerns with participant fatigue. In other words, the principal aim of the pilot study is to tune the subsequent process as much as possible.

The online survey was implemented and deployed using a tool for creating online Cloze tests. The LearnClick software tool was selected for the second pilot study (LearnClick Blog, 2016).

### **5.8.2 Why LearnClick Instrument was Chosen**

LearnClick tool is developed and deployed for use by teachers and students. It is introduced to assist teachers in assessing students' vocabulary and comprehension, and for assisting pupils to learn about context clues. However, the tool was utilised with some adjustments to suit the study requirements.

This tool is superior to other Cloze test creators. It is one of the few tools that can specialise in creating Cloze test. It has many features, flexibility, convenient and excellent for answers' assessment. This instrument has several features outlined below:

- Featuring all question types and supports the creation of Drag and Drop Cloze items that are easily used by potential participants.
- Able to create a test and store it in one place as well as the ability to store a large number of responses. Some Cloze-creation sites can only create quizzes and require you to find a place to store the responses.
- The Cloze test's questions can be presented online or printed as pdf worksheets.
- For question design, LearnClick has an option of making all blank items the same length and enables the test creator to set the number of attempts for users to submit their answers, which may reduce the false-positive rate of responses.
- The answers can be easily transferred into most common statistical formats.
- It has Statistics & Grades feature. This tool presents with extra charges (called pro membership) that presents detailed results of users' performance.
- Ease of creating quizzes. The tool's functions make it quick and easy to create tests.

On the other hand, the LearnClick has two limitations, which can be described and presented next.

- Anonymous answers cannot be recorded. In other words, each participant needs to have a unique username and a password for answers to be recorded. In order to have that, all users' emails have to be stored on the software website. This is a usability limitation of the instrument, which may affect the response rate of the study.

To minimise this obstacle, a google form has been devised as presented in Appendix B, Screenshot 5.4. So an embed link to the google form was included in the invitation email. As a result of doing this, Invitees can easily type their emails in the form. Afterwards, participants' emails will be added to the tool website. Once a participant's email stored, an automatic email will be sent that contains a unique username, a password, and the study

link. Nevertheless, this limitation could be an advantage by getting a real number of participants (as a user can only answer once).

- Chargeable Cloze-creation site (paid software). Unlike most of the available software, this quiz creator charges as a subscription fee for use. Nonetheless, this site is worth paying for using it as it offers many unique features, supportive, more powerful and flexible than other available sites.

### **5.8.3 Participants**

More than 70 LearnClick accounts have been created to people who gave preliminary approval. Nevertheless, it received only 17 completed responses (the possible causes of low participation rate will be discussed later). Each participant was randomly assigned one of two tasks. These comprised four Cloze tests and six multiple-choice questions (see 5.8.5 for procedure detail). The rest of the responses were later withdrawn because of missing answers. The remaining data of 17 subjects were entered into the SPSS program for analysis.

### **5.8.4 Materials**

Reading passages were selected from the same as the passages of the first pilot study with some enhancement and improvement. These were the set of ISPs (from native English-speaking countries such as Australia, Canada, United Kingdom and the United States). Policies were selected from a variety of countries to accommodate the possibility that the local forms of English would affect the ease of reading.

As noted earlier, this work adopted rational deletion (rational Cloze) rather than fixed ratio deletion as the main concern was to focus on 'key' components in the meaning of the sentences. In response to the respondents' feedback from the preliminary pilot study, the number of items for the Cloze test was set at 10 (20 items were for the first pre-test), one item per ISP's statement. Overall, the reading material contained 80 items (10 blanks for each of 8 ISPs). After completion of the full set of tests, all of the filled gaps were analysed and examined.

### **5.8.5 Procedures**

Prior to the main pilot study, a preliminary study was performed with 12 participants, and following analysis of the outcomes, this allowed for refining the structure of the questions and the tests subsequently adopted for the pilot study.

The procedure of the current pilot study was not that different from the first pilot study. All the selected ISPs were adopted for eight Cloze tests. The names of companies and universities were removed from the survey to prevent any unconscious bias. Since it was impractical for each participant to do all the eight Cloze tests, the eight ISPs were split into two forms. Each form was divided into three sections: starting with instruction and general information about the study, then six demographic questions (multiple-choice questions with single answers), and finally four Cloze tests (implemented as Drag & Drop interactions). The estimate for completing a form was 25 minutes, and each participant was randomly allocated to one of the two versions. After all of the members had completed, all of the responses were marked and analysed by the questionnaire designer.

One of the main purposes of performing a pilot study prior to embarking on the actual study was to test the measurement instrument (a survey). In order to make certain that the survey items perfectly covered the research questions. The pre-test attempted to realise if the survey was understandable and suitable and that all questions were unambiguous and introduced in a consistent way. The consent form, questionnaire's instructions, and Cloze test guidelines were also verified for comprehension.

### **5.8.6 Data Collection**

Various methods have been adopted for distributing the pilot study survey, including email correspondence, posting on social media (e.g., Facebook, Twitter), messaging on cross-platform messaging apps (e.g., WhatsApp, Line) and distributing flyers (contains survey URL and QR Code). Moreover, the purpose and details of the study were highlighted in the invitation letter. The invitation detailed methods of communication with the questionnaire designer. In addition, it stated clearly that

participation was voluntary, with no obligation to take part. A participant could withdraw from part or all of the study at any time without consequence.

### **5.8.7 Data Analysis**

When all the responses were received, they were first sorted into two categories according to the type of the task, i.e. G1 for the task containing policies 1 to 4 and G2 for the task containing policies 5 to 8. The respondents' gender, age, academic qualification, computer experience and study subject were also captured. All the answers were marked and verified at least twice prior to loading them onto the database sheets. Then the result analysed using the SPSS/PC statistical package. The analysis was primarily descriptive (e.g. means, median, and variance), as there were no study hypotheses to be examined. However, a number of noteworthy correlations were highlighted. This survey research did not include a hypothesis because the main aim of the research is to evaluate the ease of reading of each chosen policy as indicated by human aspect not for testing the cause and effect between survey items (Terrell, 2016).

### **5.8.8 Implications for the Main Study**

At the beginning, the survey included ten questions, six about demographics and four Cloze tests. The refinement procedures of face validity (e.g. no obvious typographical or syntactical errors) and pre-testing decreased the questionnaire to eight questions for the first pilot study, four for each of demographic question and Cloze test question, whereas for the second pre-test, six demographic questions, and four Cloze test questions. Afterwards, the final version of the questionnaire; which for the full study; contains ten questions, eight demographics (seven parent questions and one sub-question), and two questions for Cloze test. The survey for the full study has more questions of demographics than the pilot studies, in response to feedback from the pilot groups. The intent in adding two extra questions was to distinguish between respondents in terms of English language experience and nationality.

Issues that were noticed amongst respondents, included, but were not limited to, A) Understanding of the consent form, questionnaire's instructions and Cloze test directions. B) Comprehension the survey items, the terms adopted, the order of questions and the statements flow. C) The length of the survey (i.e., observing the time



spent on the whole questionnaire). D) Ability to acclimate with the survey interface (i.e., LearnClick website). E) To what extent LearnClick instrument feasibility, usability, and learnability. F) The survey design, format, and layout.

After completing two pilot studies, we can propose some suggestions for the full study as follows:

1. *The efficiency, usability and learnability of the survey software tool.*

There were many criticisms about the LearnClick software. It appeared that users require clicking twice for each question to proceed (one for submitting an answer and another for moving to the next question) and conspicuous delays after the command input to interface responding. Moreover, some users reported that they faced usability difficulty while using the website as the LearnClick website is not a mobile-friendly website. At the top of that, there were 70 people expressed an interest in the questionnaire (so a LearnClick account has been created for them), but there were only 17 complete responses. These flaws discouraged many subjects from participating.

2. *Recording responses in LearnClick website*

Although the LearnClick software is the best Cloze-creation site, it cannot record anonymous answers. To be more precise, a survey designer had to enter manually a potential participant's email into the website database, thereafter, an automatic email message will be created contains a unique username and password. This was the only method for an answer to get stored online. This technical defect led to the reluctance of many participations. Due to two obvious reasons, firstly, requesting people who you do not know their email and surname to participate, raising doubts in them that their personal privacy may be misused. This analysis was derived from that there were more than 380 clicks on the forms link, but only 70 people provided their information. Secondly, some people prefer their participation to be anonymous even to the survey creator, which cannot be achieved by using the LearnClick site.

### 3. *Changing the measurement instrument software*

One of our purposes of converting the survey from a traditional paper survey into web-based survey is maximising the questionnaire response rate. After listening to the pilot study takers' feedback, the LearnClick cannot be sufficient for the full study as respondents in the average spent more than 35 minutes to finish the survey. There is a positive correlation between a surveyed length and a participation rate, which means that our current questionnaire raises a big concern of not receiving a large number of responses. Consequently, another online survey service provider (Qualtrics, 2017) was proposed for the main study.

### 4. *Inserting further questions and features for the questionnaire*

A few demographic questions have been added or rephrased to cover the research questions. Furthermore, employing many characteristics that Qualtrics owns, for example, Survey Flow and Block Randomisation features. These features enabled the survey designer to create four blocks for the eight selected ISPs (two policies per block). By doing this, the four blocks would be evenly distributed to the survey participants so there would be a control to ensure that the blocks would be seen an equal number of times.

### 5. *Reduce the burden on prospective participants*

Survey participants would no longer require completing four Cloze tests per questionnaire as the survey had been modified to two Cloze tests. For the reason of increasing response rate.

As it can be seen from points 1, 2 and 3 there are some vital limitations of the LearnClick website, which therefore cannot be used as a measurement instrument for the larger study. The result of the pilot study survey and the software readability formula results are addressed in the following section.

## **5.8.9 Results**

Here, the research findings are presented from two perspectives, the pilot study and the readability metric.

### **5.8.9.1 Pilot Study Results**

#### **For Demographics Part**

In total, there were 17 participants in the pilot study and they were mostly males (n=15). The largest number of respondents coming from the departments of computing and mathematical sciences, and engineering and robotics, (three students from each of these two departments). For educational qualification, most of the received responses were from master's degree holders (n=12). In addition, there were two doctoral (a female and a male) and three undergraduates males. In terms of participant age, 76.5 % of respondents were between the ages of 26 to 34. The participants' experience of computer varied between proficient (n=7), intermediate (n=8), and basic (n=2). The 'below basic' option was not selected by any of the pilot study participants. Fuller demographic details of the participants are given in Table A5.1 (Appendix B).

#### **For Cloze Test Part**

All participants were provided with a task bundle consisting of instructions, guidelines and four comprehension tests. Asking potential participants to answer four Cloze tests, half the total, seemed more practical and considerably less demanding than having each of them address all eight tests. The respondents were randomly assigned to one of two groups. The first group received policies A, B, C and D (Form 1), the second group received policies E, F, G and H (Form 2), and each policy had 10 statements. The Screenshot 5.5 (Appendix B) shows one example of the Cloze tests used in this survey.

Although participants were assigned at random to complete either Form 1 or Form 2, we were also interested to determine whether there were any differences in their results of comprehension test performance between respondents addressing Form 1 and those addressing Form 2. Nine of the subjects were enrolled in the first group and their mean score was 55.55%, whereas, eight individuals in the second group had a

higher mean score (67.8%). This may be influenced by the fact that the first group included two respondents with only basic experience in computing. A further effect may have derived from the fact that the second form included two telecom organisation policies, while the first form contained only one.

With regard to performance result by participants' qualifications, the mean scores for correct answers for PhD, MSc and BSc were 65%, 61.87% and 30%, respectively. This suggests that people with higher qualification could perform better. The result appears fairly normal, in contrast to respondent computer experience' outcomes. The correct responses mean score for seven advanced level participants was 64.2 %, the correct answers mean score for eight intermediate level participants was 66.56% and the correct answers mean score of two participants with basic computer knowledge was 30%. The findings indicate that respondents with basic computing experience had a poor understanding of ISPs. The overall results are shown in Table A5.2 (Appendix B), with the values converted into percentages to make the comparison easier.

Table A5.3 (Appendix B) illustrates how participants responded to each of the Cloze tests by presenting the descriptive statistics (mean, mode, standard deviation and variance) for each examined policy. The data reveal variations in the results of the comprehension tests attributable to the difficulty of the examined policies. The findings indicate that respondents on average answered correctly more than half of the Cloze tests' items, except for policy D. The results also show that participants' performance in Policy D was somewhat low (an average of slightly above four correct answers out of ten). This showed that Policy D, as represented by the considered extracts, was the hardest policy to understand for the human reader. In contrast, Policy E was the easiest to comprehend for respondents with a mean of 7.37. With regard to tests' modes, the majority of policies had a single mode with the exceptions that Policy F and G had multiple modes. The standard deviation of the eight policies started from 1.26 up to 2.69. These low SD scores were expected due to the small sample. The findings indicated that there was great diversity among the variance scores.

### 5.8.9.2 Readability Metric Results

For our experiment, we used the Strathclyde Readability Measure (SRM1), as it is intended for text samples of more than 150 words. This software was developed previously at the University of Strathclyde and features a convenient way of selecting and comparing multiple texts for readability. The overall results, shown in Table 5.1, reveal some match between human ranking and the SRM, such as Policy D, Policy C, and Policy E. The findings indicate some similarity in ranking, which the SRM application considers as ‘close rating’ for instance, Policy A, Policy B, Policy F, Policy G, and Policy H. The SRM tool did not show any significant distance from the humans’ rating of the documents (the SRM application considers as ‘significant distance’ if there were three levels difference between user rating and SRM rating). By this, we can confirm that there are some correlations between the software readability formula’s results and the human comprehension test results, which supports our view that readability has an influence on understanding ISPs.

Text	Human		SRM1	
	Mean <sup>1</sup>	Rank <sup>2</sup>	Rank <sup>3</sup>	Scale
Policy D	43.33	1	1	41.29
Policy C	55.55	2	2	43.93
Policy H	56.25	3	5	63.61
Policy A	61.11	4	3	45.93
Policy B	62.22	5	7	64.92
Policy G	70.00	6	4	46.38
Policy F	71.25	7	6	64.40
Policy E	73.75	8	8	64.94

<sup>1</sup> Converted into percentages.  
<sup>2</sup> One is the hardest text based on human perspective.  
<sup>3</sup> Infilling cells mean that the measure rated the text same as the participants’ rate, whereas, grey cells indicate that the measure was close to the participants rating.

Table 5. 1: Comparison of human results & SRM 1 results

### 5.8.10 Conclusion

User compliance with ISPs has been extensively considered as a significant contributing element in any organisational IS plan. Despite recognising this

importance, there is a lack of literature that considers the correlation between policy provision and policy compliance.

The pilot study revealed that user compliance levels could be affected by the difficulty of understanding policy documents. This might be an effect of policy designers giving insufficient consideration to producing understandable policy materials. By applying Cloze tests to estimate human reader comprehension, our results suggest that readability, as measured using a bespoke readability metric (SRM), may yield useful insight upon the likely difficulty that end-users may face in comprehending policy documents. In turn, this supports our view that attention to the form and content of policy materials may through loss of comprehension, affect policy compliance.

## **5.9. Main Study**

### **5.9.1 Aims**

The aim of performing the main study was to use a quantitative, closed-ended, questionnaire instrument to determine the efficacy of applying readability metrics as an indicator of policy comprehensibility. The survey respondents' data were then analysed to be compared with results of nine automated readability formulae.

The quantitative method (survey) was adopted for collecting the main study data as it has many advantages including 1) it is not an expensive method, especially web-based surveys, 2) it can reach larger samples, 3) it can ensure anonymity, 4) it can deal with several topics in one questionnaire, 5) it avoids respondents' time pressure and interviewer bias, and 6) it is one of the easiest methods to code closed-ended items (De Leeuw, J. Hox and Dillman, 2012; Nardi, 2015). In addition, the quantitative method results are commonly produced from a large sample size and therefore the result can be generalised to a larger population (Scandura and Williams, 2000).

### **5.9.2 Questions**

The intent of the main study was to find an answer to the first and fourth research questions, which were as follows:

RQ1: Can the readability, as gauged by software metrics, accurately reflect the comprehensibility of ISPs? If so, how?

RQ4: Do demographic variables have an impact on reading comprehension?

### **5.9.3 Sub-questions**

Prior to presenting the main study details, the research was conducted with specific questions as described below:

Q1. Is there a significant relationship between reading comprehension and gender?

Q2. Is there a significant relationship between reading comprehension and language competence?

Q3. Is there a significant relationship between reading comprehension and the number of years studying the English language?

Q4. Is there a significant relationship between reading comprehension and age?

Q5. Is there a significant relationship between reading comprehension and academic qualification?

Q6. Is there a significant relationship between reading comprehension and computer experience level?

Q7. Is there a significant relationship between reading comprehension and study subject?

Q8. Is there a significant relationship between reading comprehension and nationality?

The mean scores of human comprehension test results clearly delineate the reading comprehension. The significance level was chosen at  $P < .05$ . The above questions will be discussed and answered in the ensuing chapter, where the main study data are analysed.

### **5.9.4 Demographic Questions**

The demographics considered in the study cover academic qualification, gender, age, nationality, study subject, computer experience and English proficiency, and provide a framework for responding to the following questions:

- ❖ What is the highest level of education you have obtained?

- ❖ What is your gender?
- ❖ What is your age?
- ❖ What is your subject area?
- ❖ How would you rate your experience with computers?
- ❖ Is English your first language?
  - ❖ How long have you been learning English?
- ❖ What is your nationality?

### **5.9.5 Sample Size**

Determining the ideal size of a research sample is a challenge frequently faced by researchers. According to Bryman (2015), deciding on the suitable size of a sample is not straightforward and there is no definitive answer. Determining sample size is a complex issue since the sample size depends on a number of factors such as restrictions on time and cost and the requirement for precision. However, he states that the larger the sample size, the greater the precision and more representative it is presumed to be. This view is supported by Saunders et al. (2009), Gray (2013), David and Sutton (2004), and Sapsford and Jupp (2006) who mention that as the size of a sample increases, sampling error decreases. The bigger the sample, the more precise the findings are likely to be.

Kline et al. (1999) declare that if the size of the sample is beneath 100 then it is considered unusable, whilst that of 150 participants is considered too small. He concludes that if the sample size is 200 or less will cause the data to be unstable as well as the test to be inconclusive. Schumacker and Lomax (2012) are in favour of Kline believes as he said that the size of a sample should be 200 or greater. Loehlin (1998) suggests that the sample size is required to be around 200 members in order to be adequate.

A popular way of determining the optimum sample size is based on the Krejcie and Morgan (1970) table, and dozens of formulae were developed to calculate an adequate sample size but according to this table. Krejcie and Morgan (1970) estimate sample size according to the 95% confidence level and the relative differences in the percentage of margin of error. These figures show a description of the precision of the



findings. The table illustrates 384 people as the optimum sample size to represent a population of 1000,000 or more at a 5% margin error.

The decision of the sample size was made to be based on Krejcie and Morgan table. According to Higher Education Statistics Agency (HESA), the targeted population is rounded off to 2.3 million higher education students in 2015/16 before the sample estimation procedure kicks in (HESA, 2017). Whereas, the total amount of higher education students remained at 2,317,880 in 2016/17, an addition of 2% from the previous statistics (HESA, 2018). Thus, 384 participants should be included in sampling at 5% margin error. It was agreed to add 216 people in the sample in anticipation of incomplete or invalid responses. A total of 600 participations were settled at for use in our research sample.

To sum up, from views introduced in the literature, the sample size for the present study should be at least 384 respondents. Since we were able to enrol 600, this seemed to ensure the results.

#### **5.9.6 Participants**

The participants for this study were local students and international students in the United Kingdom. The total number of survey responses was six hundred. One hundred and ninety-nine respondents only partially completed the questionnaire. Four hundred and one participants answered all survey questions but five respondents gave a wrong answer to the 'red herring' question. The red herring question or "trap" question is a simple question with an obvious answer, which is one of the methods to validate survey responses and discover respondents who were not engaged with the questionnaire. For survey validity, the incomplete survey and participants with the wrong answer for the red herring question were withdrawn. The remaining respondents (three hundred and ninety-six) were used for data analysis.

Participants (international students in the United Kingdom) were from a wide variety of countries including countries with English as an official language and the language of instruction in higher education. These included Australia, Canada, Kenya, Malta, Nigeria, South Africa, Tanzania, and the United States of America.

Participation in the study was on a voluntary basis. Each participant enrolled randomly in one of the four blocks to answer two Cloze tests (see procedure section for more details). The completed surveys of three hundred and ninety-six subjects were entered into the SPSS program for analysis.

### **5.9.7 Materials**

The study utilised eight out of a considered set of thirty-five policies as material for the research. These policies were a mix of public and private sectors (academia and industry). The materials were carefully selected based on a number of factors. Firstly, policies had to be from countries where English is the mother tongue. Next, the policies should be also from a variety of geographical locations, accessible online and with similar word count (no more than 10% difference). Finally, University policies were chosen from top universities of 2015 - according to QS website.

Since there were eight policies, it was a challenging task to distribute text materials to potential participants. There were two options to distribute the materials. The first option was to have participants answer questions on the eight documents. On the one hand, this option does not need many respondents and requires the shortest time for questionnaire process. For instance, this would gain 100 data for each policy from 100 participants. On the other hand, due to the length of time required for each participant to answer the whole survey for eight policies, this option would dramatically affect the willingness of participants to read the text documents and subsequently answer the Cloze tests.

We considered a second option in which each participant is required to address only one of the eight text materials. This would require more respondents to participate and complete the entire survey, which in turn means more attention paid to reading and answering the Cloze test questions. This option required less time for participants to answer the whole survey, but, in comparison with the first option, required more participants overall. For example, if 50 participants only contribute to the first text document, while the second text only receives 30 responses, to acquire 100 data per text, 800 participants would be required.

To avoid these extremes, a middle ground between these two options was selected. Two text documents out of the eight were randomly allocated to each participant. The survey was devised and deployed using the Qualtrics survey platform, which has a feature of evenly distributing two of the eight ISPs texts to the survey participants so there would be a control on the allocation of text materials (i.e., each would be seen an even number of times). This option needed less time than the first option for respondents to complete the entire survey. In addition, it should positively reflect the number of participants as well as the attention paid to answering the survey. Compared with the first and second options, 400 respondents each addressed two documents to obtain 100 participations for each text.

As noted, a number of passages from the eight (8) policies were suggested by IT experts and the role of the focus group was to choose the 10 most salient statements from the points proposed by the experts. The proposed statements were given to a small number of participants in order to obtain valuable feedback prior to proceeding to the main study. The reading passages were used for assessing the survey takers' comprehension. The materials used for the main study were slightly changed from the second pilot study. The survey was finalised with some enhancement and improvement, such as adding two extra questions to distinguish between respondents in term of English language experience and nationality, changing the online survey service provider to Qualtrics to improve usability and to benefit from Qualtrics' characteristics, and increase the response rate by asking participants to answer two instead of four Cloze tests (as discussed in detail in Chapter Five, section 5.8.8).

For a number of reasons, the Cloze test was chosen for our main study. First, a number of authors suggesting using the Cloze procedure as a method of testing the readability of a text for reading comprehension (Taylor, 1953; Bormuth, 1967; Rankin and Culhane, 1969; Guillemette, 1989; Kobayashi, 2012; Richards and Schmidt, 2015). Second, a number of readability formulae have been explored in the literature as a basis for gauging the readability of written material, but invariably depend upon syntactic variables (Klare, 1985; DuBay, 2004; Anagnostou and Weir, 2006; Campbell and Weir, 2006; Crossley *et al.*, 2017). However, Cloze test ensures that judgment of readability is not determined solely on a mechanical basis (DuBay, 2004). Last, DuBay

(2004) advised in his inclusive review on readability researches, that future studies about readability should consider the limitations of readability metrics by utilising other methods together with the readability formulae. In addition, Singer and Donlan (1982), and Biddinika et al. (2016) are in favour of using Cloze test alongside the readability formulae to measure the readability of texts.

As noted earlier, this study used rational deletion (rational Cloze) rather than fixed ratio deletion as we employed a Cloze test as a means of gauging comprehension within our experimental process. An example of rational Cloze from one of the chosen policies is shown below.

*Users must keep passwords confidential and not \_\_\_\_\_ them to others.*

In this study, words chosen to be ‘missing words’ were always considered significant in their contribution to the meaning of the phrase in which they appeared. This was the basis for the rational decision approach adopted in creating our Cloze items.

There are two types of scoring Cloze test performance; the first type is called the exact word method. The score will be counted if only the test taker writes the exact name that was deleted from the original text. The second one is known as acceptable word method (or the appropriate word method, the acceptable alternative method, or the contextually appropriate method). In this type, the score will be counted when responses are suitable or acceptable in the context. (Kobayashi, 2009; Richards and Schmidt, 2015). For the Cloze test performance in this study, the test designer preferred the acceptable word method, not the exact word method because the main aim of doing the test is to gauge the readability of texts not the language proficiency of test takers. Besides, there are native and non-native speakers participate in this test. Therefore, the Cloze test score of texts is evaluated by the percentage of words correctly entered. The lower the score, the more difficult the text.

After careful consideration, the number of items for the Cloze test with rational deletion was set at 10, one item per ISP’s statement. Overall, the reading material contained 80 items (10 blanks for each of 8 ISPs).

### 5.9.8 Procedures

Prior to the main study, two pilot studies were performed and analysed. This allowed for identifying the key elements that should be addressed in the main study. The main survey design took into consideration a number of scholars' suggestions (Bradburn, Sudman and Wansink, 2004; Scheaffer *et al.*, 2011; Dillman, Smyth and Christian, 2014) in order to increase the response rate, reduce measurement errors and obtain more accurate and usable data. The adopted suggestions were (a) utilising suitable language level, (b) avoiding obvious typographical or syntactical errors, (c) creating a simple, short and interesting survey (a well-formatted questionnaire), (d) starting the survey with a cover letter that shows an overview of the questionnaire, (e) carefully designing a survey and (f) sending a reminder notification. In addition, the researcher took into consideration the suggestions from De Leeuw et al (2012) for designing effective survey questions, including: (1) asking the right survey questions, (2) asking questions that are consistently comprehended, (3) participants can be able to retrieve information (usually from memory) required to answer survey questions, (4) questions have to be specific in order to obtain an appropriate response and (5) try to avoid as much as possible undesirable questions to get more literally accurate answers. Thereby, the researcher endeavoured to make questions straightforward, clear and easy to answer, questions can be retrieved by participants (consistently understood), and avoiding sensitive personal questions to obtain more accurate answers. In addition to that, the survey designer introduced two prizes to be awarded in order to obtain more responses as currently, many surveys are distributed and people need incentives to participate. Specifically, two prizes of gift cards were randomly drawing for those who completed the questionnaire.

A progress indicator feature was adopted in the questionnaire (see Screenshot 5.6, Appendix B). This was important for indicating respondents' progress in completing the survey. Many studies have shown that a progress indicator in a questionnaire increases the survey completion rate as well as preventing respondents from quitting, especially at the final stages (Couper, 2000; Dillman, Smyth and Christian, 2014).

The main questionnaire instrument comprised (1) a cover letter, (2) a demographic section, (3) a comprehension test section, and, in the end, (4) an option for the

respondent to provide their email, for those who wanted to be included in a prize drawn.

#### **5.9.8.1 Cover Letter**

The cover letter contained (a) a statement informing survey takers, the study purpose and the nature of the research, (b) a statement informing the participants what they have to do and the expected time to complete the self-administered questionnaire, (c) a statement informing the potential participants that participation was voluntary, (d) a statement informing questionnaire takers that they could withdraw from part or all of the study at any time without consequence, (e) a statement informing the survey taker that the study has been approved by the ethics committee of the computer and information sciences department at the University of Strathclyde (with the reference number), (f) a statement describing the prospective benefits of completing the survey, (g) methods of communication with the questionnaire designer, and (h) a thank you note for contributing to the questionnaire (see Screenshot 5.7, Appendix B). Participants signified their consent by responding to the cover letter.

#### **5.9.8.2 Demographic Part**

This part of the main study was not that different from the second pilot study. The survey questions were numbered and divided into sections, following the Nardi (2015) suggestions for clear and better formatting. In addition, the survey designer acted in accordance with Nardi advice in limiting the number of items in the demographic part to minimise survey length.

The questionnaire comprised eight demographic questions and one red herring question whereas it was six demographic questions for the second pilot study. The additional demographic questions were added for the main study to distinguish between respondents in terms of English language experience and nationality. Each survey question obtained an equal number of responses. The details of responses are displayed in frequency tables in our data analysis section.

The red herring question was added to validate survey responses, thus, any participation that containing a wrong answer to the red herring question would be discarded from the analysis procedure (Nardi, 2015). Therefore, adopting this kind of

question could help in reducing errors in the survey. The red herring question or “trap” question was a simple question with an obvious answer, which was (What is the currency in the United Kingdom?) with three options given (Dollar, Pound, or Franc). By asking this question, the aim was to ensure more quality respondents and discover respondents who were not engaged with the questionnaire. According to Scheaffe et al. (2011), questionnaire responses often include some errors, however, careful consideration in designing survey questions can minimise these errors to a point at which the outcomes are still valuable.

The questions were seeking for participant’s details on academic qualification, gender, age, nationality, study subject, computer experience and English proficiency. Nardi (2015) suggests that any question about gender, age, education, and occupation should be in the demographic section. The demographic questions from the survey are set out in Table 5.2.

<b>Demographic (Population)</b>	<b>Survey Question</b>	<b>Question Number</b>
Education	What is the highest level of education you have obtained?	1,4
	What is your subject area?	
Gender	What is your gender?	2
Age	What is your age?	3
Computer Experience	How would you rate your experience with computers?	5
Red Herring Question	What is the currency in the United Kingdom?	6
Language	Is English your first language?	7,8
	How long have you been learning English?	
Nationality	What is your nationality?	9

Table 5. 2: The main study’s demographic questions

To reveal the respondents’ comprehension of ISPs, this study compared the outcomes of the human comprehension test (Cloze test) results across the groups of education (secondary school, undergraduate degree, master degree, doctoral degree), gender (male, female), age (under 17, 17-25, 26-34, 35-54, 55 or over), computer experience (proficient, intermediate, basic, below basic), English proficiency (less 1

year, 1-3 years, 3-5 years, 5-7 years, more than 7 years), and nationality (British, Chinese, Indian, Malaysian, Nigerian, Saudi).

### **5.9.8.3 Comprehension Test**

The human comprehension test section consisted of (a) instruction on how to answer the Cloze test and (b) eight Cloze tests to gauge the respondents' reading comprehension. The Cloze test's contents were derived from the selected eight ISPs. The universities and companies' names were removed from comprehension tests to prevent any unconscious bias. Since it was impractical for each survey taker to answer all eight Cloze tests, the eight comprehension tests were divided into four blocks so each participant was randomly allocated to one of the four blocks. A block contained two Cloze tests and each one included ten statements. Therefore, the total score is 10 for each Cloze test. In general, the reading material contained 80 items, 10 blanks for each of the eight ISPs. Although there was no time limit set for survey completion, according to the online survey software tool (Qualtrics), the questionnaire could be completed in 10 minutes. After receiving a sufficient number of respondents, all of the responses were marked and analysed by the survey designer.

These comprehension tests gauged the participants' reading comprehension, which would be compared later with readability formula results in order to address whether the readability has an impact on understanding ISPs. The mean scores of human comprehension test results are indicative of reading comprehension. The significance level was chosen at  $P < .05$  as Bartlett et al. (2001) point out that a p-value of 5% is acceptable for many kinds of research.

### **5.9.9 Measures**

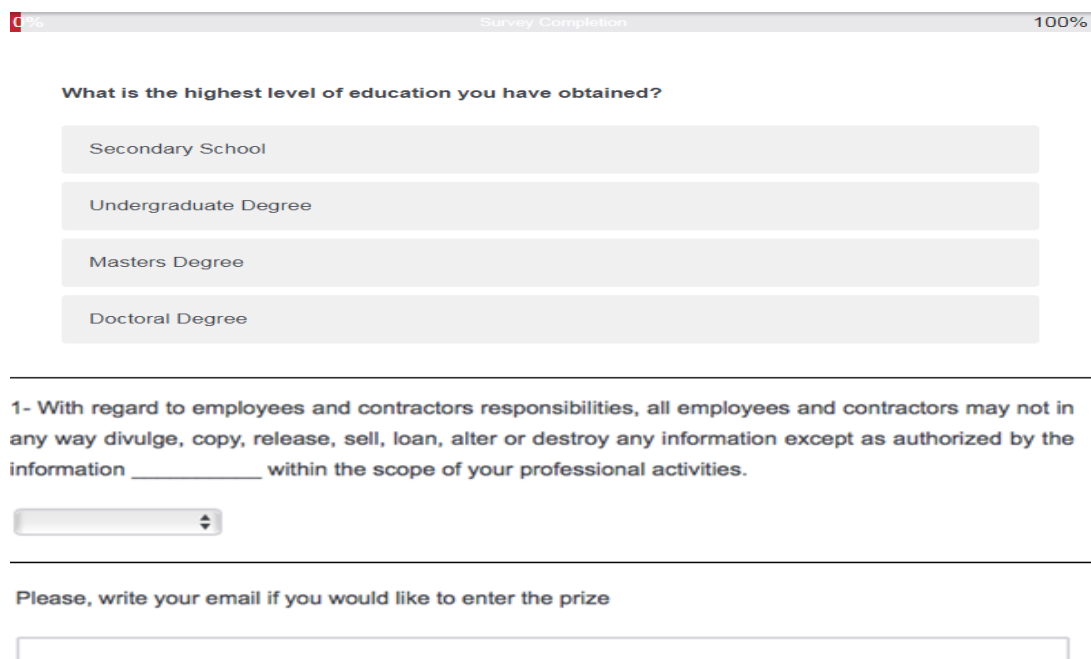
The questionnaire instrument included three parts having the following formats for gathering and gauging data on the respondents:

- Demographic part used multiple-choice format with single answers.
- Comprehension test part used dropdown list format, which had ten choices per statement with randomising choices order and single answers.
- Optional question part used text entry format with single line answer.



To make it easier to answer the demographic questions, the questions contained all likely choices for each of academic qualification, gender, age, computer experience and English proficiency data. For the question of nationality, there were a number of choices in addition to “other” choice as it was not practical to list all the nationality. For the study subject question, the designer of the questionnaire made sure to include the category “other” in the options provided in case the researcher did not anticipate all possible subjects. The category “other” had an optional choice of entering text (blank line) for who would like to write more details.

To mitigate learning effects, the choices order of comprehension test statements were randomised with a feature of dropdown list format. Example question of the three sections is shown in Figure 5.4.



The screenshot shows a survey completion interface. At the top, there is a progress bar labeled 'Survey Completion' with '100%' on the right. Below this, a question is displayed: 'What is the highest level of education you have obtained?'. A dropdown menu is open, showing four options: 'Secondary School', 'Undergraduate Degree', 'Masters Degree', and 'Doctoral Degree'. Below the question, there is a legal disclaimer: '1- With regard to employees and contractors responsibilities, all employees and contractors may not in any way divulge, copy, release, sell, loan, alter or destroy any information except as authorized by the information \_\_\_\_\_ within the scope of your professional activities.' Below the disclaimer is a small dropdown menu with a downward arrow. At the bottom, there is a text prompt: 'Please, write your email if you would like to enter the prize' followed by an empty text input field.

Figure 5. 4: Example of demographic question types

### 5.9.10 Data Collection

The main study survey was devised and deployed using the Qualtrics survey platform. Despite the fact that the Qualtrics software tool was not designed specifically for creating online Cloze tests, Qualtrics was adopted because it could meet our requirements for creating the Cloze test question structure and had many features, such as sample management, user-friendly interface, storage of data, and options for survey

layout and question forms. In addition, the Qualtrics platform has other useful and unique features (e.g. Survey Flow and Block Randomisation features). The survey designer was able to create four blocks for the eight selected ISPs (two policies per block) and ensure the four blocks would be evenly distributed to the survey participants so there would be a control to ensure that the blocks would be seen an equal number of times. Qualtrics exports participants' answers into most commonly used statistical formats such as the SPSS.

Various methods have been adopted for distributing the online survey, including email correspondence (including mass mailing to all Strathclyde international students), posting on social media (e.g., Facebook, Twitter), messaging on cross-platform messaging apps (e.g., WhatsApp, Line) and distributing flyers (contains survey URL and QR Code).

A quantitative data analysis approach was used to analyse the survey data once the required number of participants was reached, and the completed survey data were collected and exported to SPSS. As it was important for analysis to receive completed questionnaires, any partial responses were discarded. Storing survey data in a secure environment was a top priority of the survey designer. The data were kept on the researcher's personal network drive and the data were only accessed by the authorised user. As soon as the data are no longer required, it would be destroyed comprehensively and securely. Figure 5.5 illustrates the steps taken to design and probe the survey instrument.

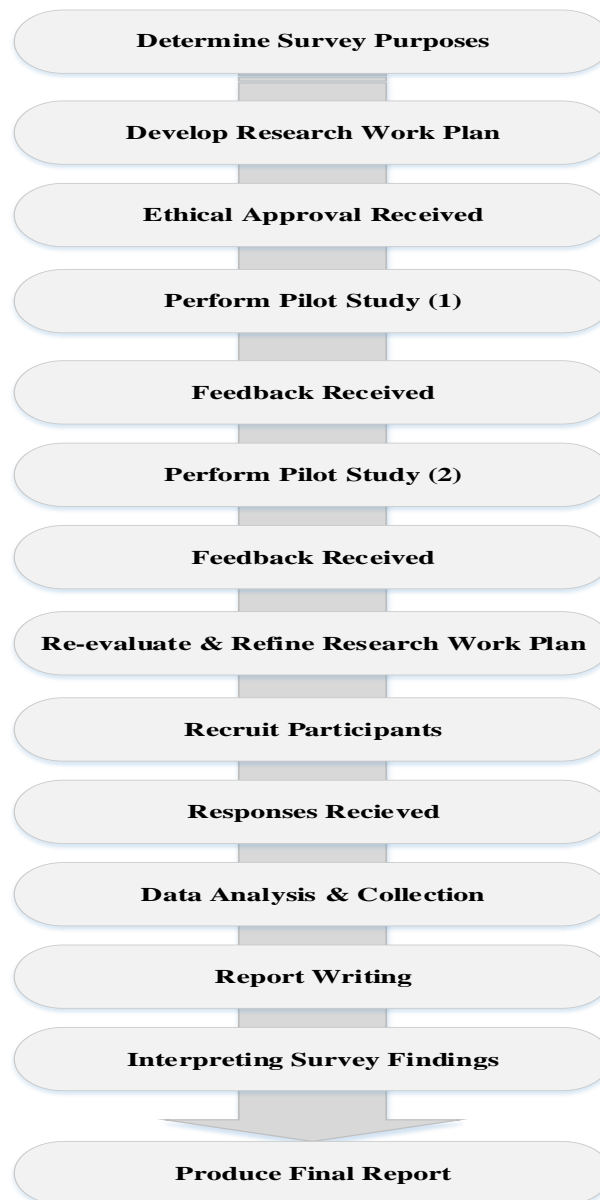


Figure 5. 5: The main study process

### **5.10.Summary**

The details of the research design, methods and data collection are identified in this chapter. In this study, we conducted two rounds of pilot tests with a number of people to ensure the questions were well-worded and understandable, to make sure the feasibility of survey and the proposed incentives, verifying the clarity of survey instructions, to ascertain the survey questions reflect the desired end of adopting the questionnaire method, the consistency and the rational sequence of the survey items.

Besides, an exploratory sequential mixed methods approach was adopted to enable us to collect qualitative data, which is used later for defining the Cloze test questions, and then these questions are used in our quantitative method (survey), on the next chapter, and the outcomes of exploratory sequential approach will be then compared against a number of readability formulae to identify the degree of similarity and difference between the results of human and software metrics.

The following chapter presents the data analysis of the main study (Phase6) and compares its results with a number of readability metrics (Phase7).

## **CHAPTER SIX: RESULTS AND DATA ANALYSIS**

### **6.1. Introduction**

This chapter presents the quantitative analysis of data sets that were collected by the investigator in the main study. For validity purpose, this research examines the outlier cases and normality on the data before the data are approved to be included in our study. The data collection and analysis methods were introduced earlier in Experimental Design Chapter. The study has received 600 survey responses and following Janssens et al., (2008) advice of using statistical analysis as it helps in sorting out and sums up large data sets in a smaller number of meaningful statistical displays. This is accomplished here by introducing, describing and discussing in detail the demographic profile of the participants' gender, language, number of years for studying the English language, age, qualification level, computer experience level, study subject, and nationality. This chapter shows the results of the eight selected information security policies (ISPs) and then compares between respondents' results with a number of readability formulae.

### **6.2. Descriptive Statistics**

The self-administered questionnaire results were collected from Qualtrics, imported and stored into the Statistical Package for Social Sciences (SPSS) program for statistical analysis purpose. As a method of presenting quantitative descriptions in a manageable form (Babbie, 2016), descriptive statistics were utilised to describe the features of the data in the study. This type of statistics highlights how 'variables are allocated' and readily reflects the responses of the subjects (Holbert, 2013). According to Gray (2013), descriptive statistics can describe the basic features of a research, usually throughout the use of graphical analysis.

The survey consisted of three sections, two of these were required and the last section was an optional question. Responses were sorted into four categories according to the type of the task answered, i.e., group1 for the task containing policies one and two, group2 for the task containing policies three and four, and so on. Each group

received an almost equal number of responses, between 94-104 responses. The respondents' academic qualification, gender, age, nationality, study subject, computer experience and English proficiency were captured. All answers were reviewed and verified at least twice before being imported into the results database for subsequent statistical analysis, using the SPSS/PC statistical package. A t-test and one-way ANOVA test were performed to determine which factors had significant effects on reading comprehension.

The result of the main study survey and the software readability formulae results are addressed in the following section.

### **6.3. Results**

In the examination of the responses received, there were a total of 600 people enrolled and as a result, a sample of 396 usable responses was collected (66% of the total participation). For survey validity, incomplete participation or complete participation but not giving the correct answer to the red herring question were discarded and not considered.

#### **6.3.1 Demographics**

Respondents for the main study were secondary education, post-secondary education and postgraduate education students, local students and international students, in the United Kingdom.

A summary of the results of the survey is now presented, utilising figures, tables and a commentary on the various findings.

##### **6.3.1.1 Response by Gender**

Based on analysis of responses from the surveyed volunteers, as presented in Figure 6.1, from the 396 contributors included in the questionnaire it was found that the split of respondents almost equally distributed by gender with 51.3% males and 48.7% females.

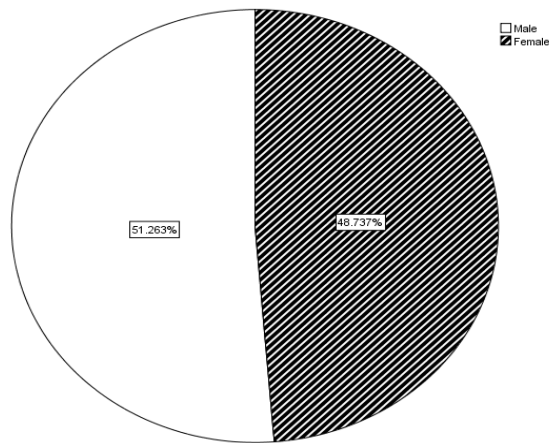


Figure 6. 1: Response by gender

### 6.3.1.2 Response by Language

Of the 396 surveyed participants, the majority of respondents did not have the English as their first language. This represented 75.8% of the sample, with the remaining 24.2% having English language as their first language, as set out in the pie chart below (Figure 6.2).

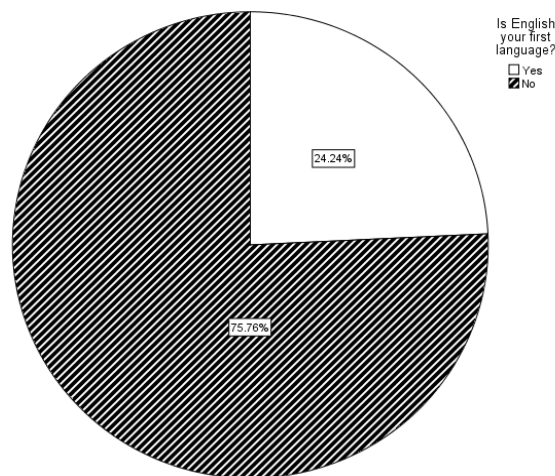


Figure 6. 2: Response by language

### 6.3.1.3 Response by Number of Years for Studying the English Language

This sub-question was asked only for non-native English speakers, in order to provide insight on how long they have been studying the English language. The pie chart below (Figure 6.3) indicates that all five categories received responses. The majority of non-native English participants (300 people) have learned English for more

than seven years, which represented 49.7% of the total sample. A smaller group of 21% had studied English for between three to five years at the time of data collection, and 16% of the participants had studied English for between one to three years. Of the remaining participants, 10% had studied English for between five to seven years, and only 3.3% of the non-native English participants said that they had studied the English language for less than one year.

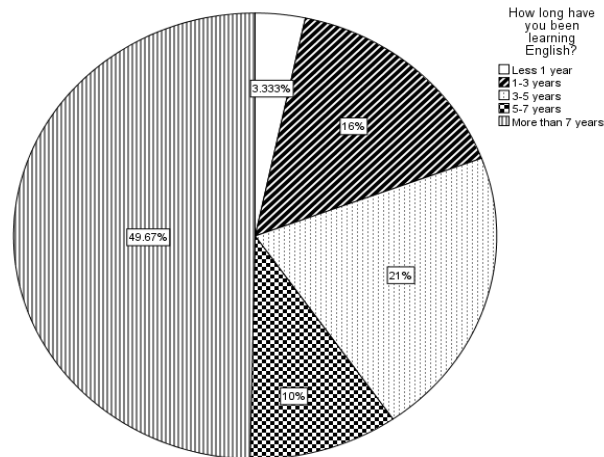


Figure 6. 3: Response by number of years for studying the English language

#### 6.3.1.4 Response by Age

From the data in Figure 6.4, it is apparent that the majority of the cohorts (49.5%) were in the 17-25 age group because this is usually the university age. The next largest group (39.9%) was in the 26-34 age group. About 10% of those surveyed were aged between 35 to 54, and only two respondents identified as being within the age group of 55 or over.

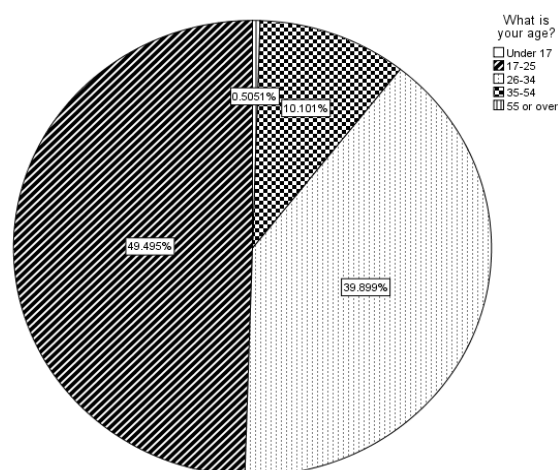


Figure 6. 4: Response by age



### 6.3.1.5 Response by Qualification Level

This demographic question asked each contributor about their highest level of education (Figure 6.5). All four categories received responses with the majority of participants having obtained a BSc degree (40.9%). In the next largest group, 33.8% of respondents were educated to at least postgraduate level (MSc graduates), while 19.7% of the participants had only secondary school education and, in the smallest group, only 5.6% of the participants – had a PhD degree.

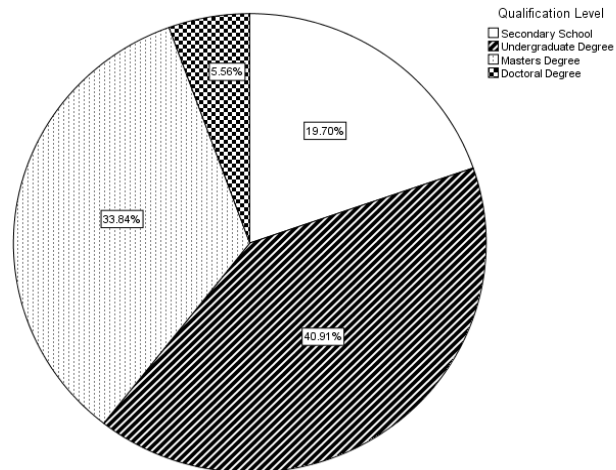


Figure 6. 5: Response by qualification level

### 6.3.1.6 Response by Computer Experience Level

Based on the analysis of all 396 surveys, as shown in the pie chart below (Figure 6.6), the majority of respondents considered their computer experience as intermediate, represented by 52.3% of the sample. 37.1% of participants evaluated their computer experience as proficient, and the remaining 10.6% of respondents rated their computer experience as basic.

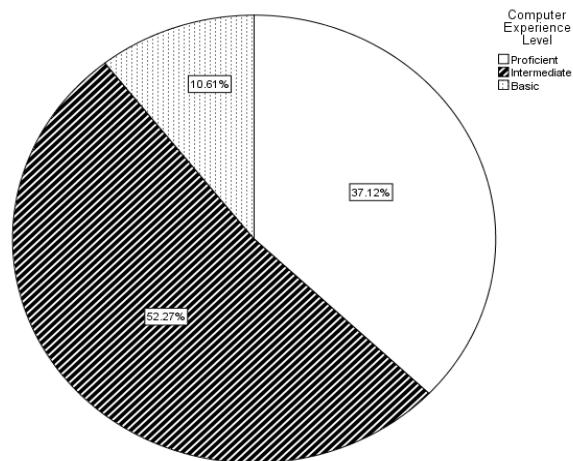


Figure 6. 6: Response by computer experience level

### 6.3.1.7 Response by Study Subject

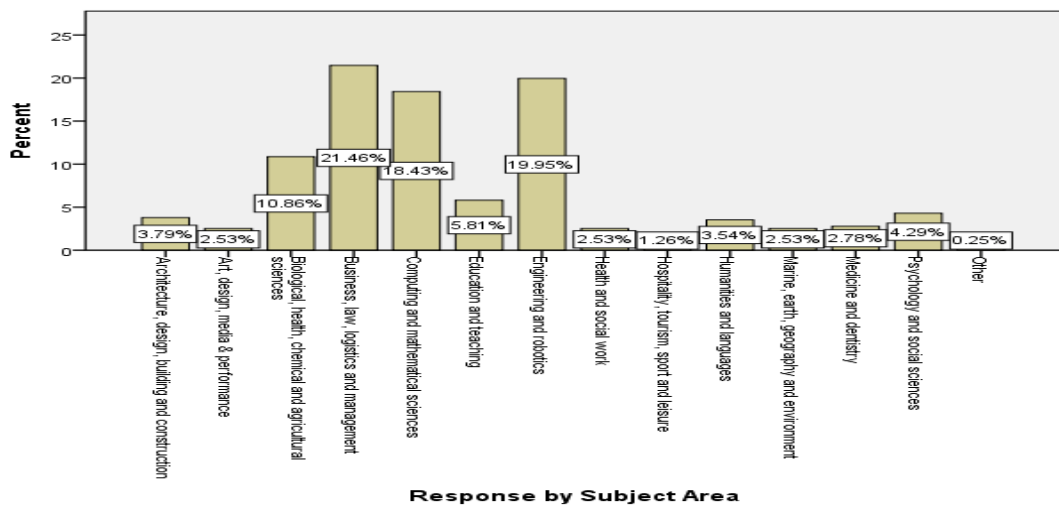


Figure 6. 7: Response by study subject

Figure 6.7 depicts the distribution of the 396 participants' educational disciplines and shows that all fourteen categories received responses. The majority of participants were from four categories, the first (21.5%) was business, law, logistics and management department, the second (19.9%) was engineering and robotics department, and the third (18.4%) was computing and mathematical sciences department. At the fourth level, 10.9% of the surveyed participants were under the category of biological, health, chemical and agricultural sciences department. Noticeably, the other categories (10) each had lower than 6% and only one respondent indicated that they were not in any of the listed subjects.

### **6.3.1.8 Response by Nationality**

The last demographic question asked the nationality of each participant. Interestingly, the 396 responses were from forty-five countries. From the data in the Table 6.2, it is apparent that most of the respondents were from Saudi Arabia (22.7%). There was no intention to have a majority of Saudi participants, as the questionnaire was spread online without targeting particular nationalities and the majority of responses happened to be from Saudi nationals. At the second level, responses from two other national groups were fairly close in number, with 12.1% of participants being British and 11.9% of participants being Chinese. Remarkably, the remaining groups had lower than 7% and 3.3% of the respondents had not specified their nationality.

The results obtained from the demographic questions are summarised in Table 6.1 and Table 6.2, below, which details the responses in frequency tables. It is worth mentioning that, in total 396 surveyed participants completely answered the questionnaire and only the data from these surveys were included in the analysis.

	Count	Percent
<b>Gender</b>		
Male	203	51.3
Female	193	48.7
<b>Language</b>		
English	96	24.2
Others	300	75.8
<b>Learning English (for non-native English speakers)</b>		
Less 1 year	10	3.3
1-3 years	48	16
3-5 years	63	21
5-7 years	30	10
More than 7 years	149	49.7
Total	300	100.0
<b>Age</b>		
17-25	196	49.5
26-34	158	39.9
35-54	40	10.1
55 or over	2	.5
<b>Qualification</b>		
PhD	22	5.6
MSc	134	33.8
BSc	162	40.9
Secondary school	78	19.7
<b>Computer Experience</b>		
Proficient	147	37.1
Intermediate	207	52.3
Basic	42	10.6
<b>Study Subject</b>		
Architecture, design, building and construction	15	3.8
Art, design, media & performance	10	2.5
Biological, health, chemical and agricultural sciences	43	10.9
Business, law, logistics and management	85	21.5
Computing and mathematical sciences	73	18.4
Education and teaching	23	5.8
Engineering and robotics	79	19.9
Health and social work	10	2.5
Hospitality, tourism, sport and leisure	5	1.3
Humanities and languages	14	3.5
Marine, earth, geography and environment	10	2.5
Medicine and dentistry	11	2.8
Psychology and social sciences	17	4.3
Others	1	0.3

Table 6. 1: The demographic information

	Count	Percent
<b>Nationality</b>		
American	15	3.8
Australian	1	.3
Austria	1	.3
Bolivian	1	.3
British	48	12.1
Bulgarian	14	3.5
Canadian	4	1.0
Chinese	47	11.9
Colombian	2	.5
Czech	3	.8
Danish	1	.3
Dutch	2	.5
Egyptian	2	.5
Emirati	1	.3
Estonian	2	.5
Finnish	1	.3
French	5	1.3
German	7	1.8
Greek	13	3.3
Honduran	1	.3
Hungarian	1	.3
Icelandic	1	.3
Indian	18	4.5
Indonesian	2	.5
Iranian	3	.8
Irish	2	.5
Italian	4	1.0
Kenyan	1	.3
Malaysian	25	6.3
Maltese	2	.5
Nigerian	19	4.8
Not specified	13	3.3
Omani	2	.5
Pakistani	6	1.5
Polish	10	2.5
Portuguese	1	.3
Romanian	3	.8
Russian	1	.3
Saudi	90	22.7
Slovakian	4	1.0
South African	1	.3
Spanish	7	1.8
Swedish	1	.3
Tanzanian	1	.3

Thai	6	1.5
Ukrainian	1	.3
Total	396	100.0

Table 6. 2: Response by nationality

### 6.3.1.9 The Mean of the Survey Elements

Table 6.3 below, presents the total average value of participants according to their answers on Cloze tests and shows the total mean of some demographic questions. In addition, the table indicates the total number of participants for each presented groups' elements in the table.

As per the table presentation, there were variations in the total numbers of responses in the four groups (despite the Qualtrics survey platform having evenly assigned potential participants to one of the four groups). This could happen if a participant did not complete the survey or they failed in answering the red herring question so their participation was withdrawn. In other words, 396 out of 600 survey responses were entirely completed and therefore only the data from these questionnaires were considered for the analysis. In terms of mean scores, the findings indicated that respondents on average answered correctly more than half of the Cloze tests' items.

In terms of the survey groups, participants' total average value within the second group obtained 10.98 out of 20, which was the lowest value among the four groups. Whereas, the fourth group of participants gained the highest mean of 15.07. The total average value for the first group and third group were 12.72 and 13.15, respectively. The relatively small variation in the mean between groups could result from difficulties in reading the texts.

For the gender aspect, the total average value for the female participants, with a mean of 13.59, was higher than the male participants, with a mean of 12.48. The result indicated that the female performance was better than the male performance in understanding security policy concepts. A number of reports (Cappelli *et al.*, 2009; Grant, 2010; Hanley *et al.*, 2011) support this finding by signifying that females pay more attention to ISPs than males. In contrast, a study by Kowalski *et al.* (2008), indicated that both genders pay equal attention to information security (IS).

In terms of participants' language, 96 (24.2%) people were the native speaker of English, whereas, 300 (75.8%) participants were non-native English speakers. An Independent Samples t-test would be applied to investigate the language competence effect on the study research results (see section 6.3.4 for more details). Noteworthy, the ISP documents are intended for both native and non-native English speakers.

From the table below, we see that native participants obtained better mean scores (Mean=15.8646) than their non-native English counterparts (Mean=12.1167). This result was expected, as people usually understand their first language better than a foreign language.

On the basis of the analysis shown in Table 6.3 of their years of receiving English language teaching, a direct relationship was identified between the comprehension of ISP texts and the number of years of learning the English language. This detail was a sub-question asked only of non-native English speakers, thereby, 300 responders were given this question. The results can be understood, as the more a person spends time in learning a language the more vocabulary they know, therefore, the more likely they will comprehend an ISP's content.

However, some results for the groups of (less than year) and (1-3 years) in learning English raised concern as they gained lower than the average of the total mean score. This indicates that they may have agreed and signed to adopt ISP documents when they understood less than half.

From the result obtained from the age question, it was found that participants who fall within the age group of 55 or over had the highest mean rank (15 out of 20). Therefore, it can be assumed that a student of this age has more knowledge and experience on ISPs than other younger students. Secondly, those aged 17-25 were in second place regarding the mean rank (M=14.05). This result might indicate that students at this age are learning or they have just learned about computing and security concepts. There was no scientific reason why the age groups of 26-34 (M=11.71) and 35-54 (13.07) gained lower mean rank than the other age cohorts.

In terms of surveyed participant's qualification, the ranks of the mean scale from the top to bottom were as follows: students of secondary school, undergraduate degree,

master degree, and then doctoral degree. The results showed that the majority of participants (162) have attained at least graduate level, 134 respondents were educated to postgraduate level (MSc) and 78 of the surveyed participants had only attended secondary school. Only 22 participants – the smallest group – had PhD level of education.

The surprising results indicate an inverse relationship between education level and understanding ISP contents, which was in contrast to some of the readability formulae principals in terms of the ability to comprehend a text based on a reader's level of education. Contrariwise, our findings correspond to the results reported by Grant (2010). He proposed that respondents with higher education level were less security aware than participants with lower levels of education.

On the basis of computer experience, it can be acknowledged that the highest mean score (14.1) was obtained by the proficient computer experience participants. Then, counterparts with an intermediate experience in computing (12.72 mean score). Next, respondents who had only a basic knowledge in the computer field (10.69 mean score). It was noticeable that the computer experience level was directly proportional to the comprehension of ISP content. Thus, the result implied that individuals with more expertise in computing could perform better. These findings echo the work of Alotaibi et al. (2016).



	<b>Participants' count</b>	<b>Mean*</b>
<b>Group</b>		
First group	97	12.7216
Second group	94	10.9894
Third group	104	13.1538
Fourth group	101	15.0792
<b>Gender</b>		
Male	203	12.4828
Female	193	13.5959
<b>Language</b>		
English	96	15.8646
Others	300	12.1167
<b>Learning English (for 300 non-native English speakers)</b>		
Less 1 year	10	5.8000
1-3 years	48	8.7917
3-5 years	63	11.6825
5-7 years	30	12.6667
More than 7 years	149	13.6846
<b>Age</b>		
17-25	196	14.0510
26-34	158	11.7152
35-54	40	13.0750
55 or over	2	15.0000
<b>Qualification</b>		
Secondary School	78	15.0128
Undergraduate Degree	162	13.0000
Master Degree	134	12.2761
Doctoral Degree	22	10.7273
<b>Computer Experience</b>		
Proficient	147	14.1088
Intermediate	207	12.7295
Basic	42	10.6905
*The mean is out of 20		

Table 6. 3: The mean of the survey elements

#### 6.3.1.10 Overview of the Examined Policies Results

The survey's policies were compiled with details that include mean, standard deviation and the number of respondents, as shown in Table 6.4. To clarify policy

scores, the minimum possible score was zero and the maximum possible score was 10. The total average value was out of 10.

It is worth mentioning that, the eight policies were divided into four blocks (by the Qualtrics survey system) and each responder was randomly allocated to one of the four blocks. A block contained two policies and for each policy included ten statements. In general, the reading material contained 80 items, 10 blanks for each of the eight ISPs. Because of this division, it can be noticed that policy A and B, policy C and D, policy E and F, policy G and H had the same number of responders (see Table 6.4).

For policy A, 97 responders answered this policy. The mean for policy A was 6.5052 with a standard deviation of 2.6501. The minimum cumulative score was zero (2 participants), while 16 participants reported the maximum possible score of 10.

For policy B, there were 97 completed answered. A mean value of 6.2165 and a standard deviation of 2.4462 were the results. The minimum score of zero was calculated for five participants, while eight participants reported a maximum possible score of 10.

For policy C, 94 people participated in, with a mean of 5.9043 and a standard deviation of 2.2049. Possible values allowed were from a minimum of zero to a maximum of 10, two responders scored the lowest possible score whereas three responders scored the highest possible score.

For policy D, there were 94 participations have received. A mean value of 5.0851 and a standard deviation of 2.6054 were the results. The minimum score of zero was calculated for three participants and six participants gave the maximum score of 10.

For policy E, 104 participations included in. The mean for policy E was 7.0962 with a standard deviation of 2.4947. Two responders scored zero, which was the lowest score, and the maximum score achieved was 10 by 17 participants.

For policy F, 104 responders answered this policy. The mean for policy F was 6.0577 with a standard deviation of 2.8483. Eleven responders scored the highest with a score of 10, and 4 responders scored zero, which was the lowest of the group.

For policy G, there were 101 completed answered. A mean value of 7.9208 and a standard deviation of 2.5755 were the results. The minimum score of zero was calculated for only two participants, while the maximum score achieved was 10 by 47 participants (where the largest number of responders gained the highest score among the 8 policies).

For policy H, 101 people participated in, with a mean of 7.1584 and a standard deviation of 2.4443. Possible values allowed were from a minimum of zero to a maximum of 10, three responders scored the lowest possible score whereas 23 responders scored the highest possible score.

	<b>Policy A</b>	<b>Policy B</b>	<b>Policy C</b>	<b>Policy D</b>	<b>Policy E</b>	<b>Policy F</b>	<b>Policy G</b>	<b>Policy H</b>
<b>Participants' count</b>	97	97	94	94	104	104	101	101
<b>Mean*</b>	6.5052	6.2165	5.9043	5.0851	7.0962	6.0577	7.9208	7.1584
<b>Std. Deviation</b>	2.6501	2.4462	2.2049	2.6054	2.4947	2.8483	2.5755	2.4443
*The mean is out of 10								

Table 6. 4: Overview of the examined policies results

All in all, the insight is drawn from Table 6.4 was that the responder numbers ranged from 94 to 104. It is apparent from this table that a wide range of mean value in the examined ISPs, beginning with 5.0581 and ending with 7.9208. The standard deviation range between policies was convergent, starting from 2.2049 up to 2.6501. All policies had responders with minimum and maximum possible score, but policy B had the largest number of responders with a minimum score (by five in numbers), whereas policy G had the largest number of responders with a maximum possible score by 47 participations.

Notably, the results showed that participants' performance in policy D and policy C were somewhat low (an average of slightly less than six correct answers out of ten). This showed that policy D, as represented by the considered extracts, was the hardest policy to understand for the human readers (which matched the second pilot study outcomes). In contrast, policy G was the easiest to comprehend for respondents.

### **6.3.2 Participants' Results VS Traditional Readability Index's Outcomes**

This study selected eight of the most common readability formulae to be compared with human results of Cloze tests. Taking into consideration DuBay's (2004) suggestion of using other methods in conjunction with readability formulae, the eight selected policies were ranked twice, based on participants' performance of Cloze tests, and based on a software readability formula perspective (as set out in Table 6.5).

The eight chosen readability metrics are as follows, Flesch Reading Ease (FRE), Flesch-Kincaid Grade Level (FKGL), Gunning Fog Index (FOG), Spache, "new" Dale-Chall, SMOG, Coleman-Liau and Automated Readability Index. The scores of the eight formulae were obtained from a website called readable.io. Among the websites that offered readability scores, this was the only one that supported all the eight formulae. It is worth mentioning that the Spache readability formula was discarded as it performs best on text documents that are for children up to fourth grade, while our selected ISPs documents were written for people who are older than elementary level of education. Notably, the analysis of ISPs was not based on some samples of text, but on the entire text document that was accessible to be analysed.

The ordinal structure of the rank is simple and clear, as one (1) indicates the hardest text in the set and eight (8) indicates the easiest text in that set of texts. To clarify Table 6.5, there are three types of cell. If the cells of 'Rank' columns are shaded, this indicates that the measure was close to the human rating. Unshaded cells mean that the measure rated the text difficulty as the same as the participants' rating and triple lined borders on a cell indicates a significant distance from the human rating of the text.

In terms of the FRE measurement category, all examined policies' texts were considered as very difficult to read (college graduate level) except for two texts considered as difficult to read (college level). In terms of the comparison with the human ranking, the majority of text documents' results (six out of eight policies) were close to the human ranking with two texts indicated as a significant distance from the human rating of the text.

For the "new" Dale-Chall tool, the surveyed documents were categorised variously between seventh or eighth grade, ninth or tenth grade, and eleventh or twelfth grade

(based on the United States education system). In comparison with the human ranking, all policy document results were close to the participants' rating except for one document, which indicated a significant distance from the human rating of the text.

In terms of the FOG index, all of the text documents were measured as too complex for most humans to read. The reading levels ranged from college junior to postgraduate level. In terms of comparison with human ranking, four ISP texts registered as close to the respondents' ranking, two documents were identical to the human ranking and two policies emerged with a significant distance from their human ratings.

For the SMOG index tool, all of the examined ISPs texts were considered as very difficult to read. The reading levels of the documents were categorised as college junior up to postgraduate level. In comparison with the human ranking, the majority of document results (five out of eight policies) were close to the human ranking, two texts indicated a significant distance from the human rating of the text, and for one document, the participants and SMOG tool gave the same rating.

For the Coleman-Liau formula, the reading level measurement varied between high school senior up to college graduate. In terms of the comparison with human ranking, half of the chosen policies were close to the human rating. In addition, two of the text documents received the same rating as the respondents' rating while two policies showed a significant distance from their human rating.

For the Automated Readability Index, the surveyed documents were rated between twelfth grades up to postgraduate level. In comparison with the human ranking, half of the examined policies' texts had the same rate as the participants' rating. Moreover, three of the text documents were close to the participants' rating, whereas, only one document indicated a significant distance from the human rating for the text.

All in all, the results across the different measures revealed that none of the selected traditional formulae give identical scores nor do they match the human rating. The differences in rating the same piece of text are not surprising given previous studies (Klare, 2000; Zhou, Jeong and Green, 2017). Furthermore, the analysis showed that most of the surveyed policies considered as fairly difficult to read by traditional readability formulae were also found difficult by native English readers. It is worth

keeping in mind that these ISPs documents are intended for both native and non-native English speakers. Therefore, we speculate that these results were due to the many policy designers do not consider ease of comprehension when they write ISP documents.

	Human Rank	FRE	Rank	New DC	Rank	FKGL	Rank	Fog	Rank	SMOG	Rank	Coleman Liau	Rank	Automated Index	Rank
Policy D	1	17.5	4 <sup>c</sup>	8.1	4	18.7	4	19.4	4	18.9	4	14.4	7	19.5	3 <sup>d</sup>
Policy C	2	8.3	3 <sup>a</sup>	8.9	1 <sup>d</sup>	19.1	3	20.9	1 <sup>d</sup>	19.4	3	16.8	3	19.7	2
Policy F	3	0.3	1	8.9	1 <sup>d</sup>	19.8	1	20.1	3 <sup>b</sup>	20.1	1 <sup>d</sup>	17.7	1	19.5	3 <sup>d</sup>
Policy B	4	4.6	2	8.7	3	19.7	2	20.9	1 <sup>d</sup>	20.1	1 <sup>d</sup>	17.6	2	20.4	1
Policy A	5	27.9	6	6.8	7 <sup>d</sup>	15	5	17.8	5	16.7	5	14.6	5	15.3	5
Policy E	6	30.6	7	6.8	7 <sup>d</sup>	13.4	8	15.9	7	15.7	7	14.5	6	13.1	8
Policy H	7	36.3	8	7.1	5	13.7	7	15.0	8	14.8	8	12.6	8	13.4	7
Policy G	8	27.5	5	6.9	6	14	6	16.8	6	16.1	6	14.7	4	13.5	6

- a. Shaded cell indicates that the measure was close to the participants' rating.
- b. Unshaded cell means that the measure ranked the text same as the participants' rating.
- c. Triple lined borders on a cell indicate a significant distance from the human rating of the text.
- d. The policy ranking is same as another policy in the same measurement tool.

Table 6. 5: Comparison of human results VS traditional formulae results

### 6.3.3 Participants' Cloze Test Results VS Strathclyde Readability Measure1 Outcomes

As discussed earlier, survey contributors were assigned two Cloze tests and the content of the tests was taken from two policies. This study examines and compares eight ISP text documents on eight mechanical readability formulae's results with a human-based comprehension test outcomes. The last formula to be examined and compared with our selections of text documents is the new measure of readability formula (Strathclyde Readability Measure1).

Here, the rank indicates the relative text difficulty, with a rank of one (1) indicating the hardest text and a rank of eight (8) indicating the easiest text (see Table 6.6). The cells of 'Rank' columns are shaded or unshaded. Unshaded cells mean that the readability measure ranked the text the same as the participants' rate, whereas, shaded cells indicate that the measure was close to the human rating. Notably, the policies' mean value was presented out of 10; while, the maximum scale of policies were 100.

In addition, the analysis of ISPs was not based on ISP text extracts, but on the entire text documents.

Text	Human Mean	Rank		SRM1 Scale
		Human	SRM1	
Policy D	5.0851	1	1	40.44
Policy C	5.9043	2	2	42.65
Policy F	6.0577	3	3	43.39
Policy B	6.2165	4	4	43.77
Policy A	6.5052	5	5	46.15
Policy E	7.0962	6	7	62.46
Policy H	7.1584	7	6	61.63
Policy G	7.9208	8	8	63.77

Table 6. 6: Comparison of human results VS SRM 1 results

Considering the Strathclyde Readability Measure (SRM) results, two-thirds of the surveyed policies were considered fairly difficult. We speculate that these results suggest that many policy designers do not consider ease of comprehension when writing ISP documents.

The results, summarised in Table 6.6, reveal remarkable findings. The majority of texts' results were identical in human ranking and the SRM. In other words, six out of eight policies gained the same rating in both measurements. In addition, the findings indicated some similarity in ranking, which the SRM application considers as 'close rating', as can be seen for Policy E and Policy H. As a side note, Table A6.1 (Appendix C) shows the eight policies' name.

The SRM tool did not show any significant distance from the humans rating of the documents (the SRM application considers as 'significant distance' if there were three levels difference between user rating and SRM rating). Thereby, we confirm that there were correlations between this software readability formulas' results and human comprehension test results, and this strongly supports our view that the readability influences the understanding of ISPs. In addition, difficulty in reading such documents may contribute to user non-compliance with ISPs.

Furthermore, these results suggest that an application such as SRM may be used in evaluating the readability of a text as it differs in approach from most other readability measures. This is the closest approach we have found to estimate human reader

comprehension and affords a practical alternative to Cloze-based comprehension tests. Aside from this, all policies gained more than 5 out of 10 in mean score, which means that the respondents (N=396) on average correctly answered more than half of the Cloze tests' items.

#### **6.3.4 Key Findings of Demographic Variables**

This subsection considers the influence of demographic variables (gender, language, number of years for studying the English language, age, qualification level, computer experience level, study subject and nationality) on perceived reading comprehension. In this study, there were 396 respondents with 203 males (51 percent) and 193 females (49 percent). In the sample, there were 300 non-native English speakers (76 percent) and 96 native English speakers (24 percent). It is within the non-native English speakers group; participants were classified into five categories based on their years of learning English. These categories were as follows: 1) less than a year comprising only 3.3 percent of the sample; 2) 1-3 years comprising 16 percent of the total sample; 3) 3-5 years comprising 21 percent of our sample; 4) 5-7 years comprising 10 percent of the sample; and 5) more than 7 years comprising nearly 50 percent of the entire sample, which represented the majority of cohorts.

In term of age, there were four groups. These groups were as follows: 1) 17-25 years comprising approximately 50 percent of the total sample, which represented the majority of participants; 2) 26-34 years comprising 40 percent of the sample; 3) 35-54 years comprising 10 percent of the entire sample; and, in the smallest group, 4) from 55 years and above comprising less than 1 percent of the selected sample. There were four main levels of education groups. These groups were secondary school (nearly 20 percent), undergraduate (closely 41 percent) (the highest percentage of the sample), master's degree (approximately 34), and doctoral degree (5 percent). In term of computer experience, there were three categories. These categories were proficient (37 percent), intermediate (52 percent), and basic (approximately 11 percent).

Furthermore, in the sample, there were fourteen groups based on their study subject. The highest percentages of the sample were four groups as follows: 1) business, law, logistics and management department comprising nearly 22 percent of the total sample; 2) engineering and robotics department comprising 20 percent of the



sample; 3) computing and mathematical sciences department comprising 18 percent of the entire sample; and 4) biological, health, chemical and agricultural sciences department comprising 11 percent of our sample. Whereas other groups were each one of them below 10 percent of the total sample. In term of nationality, there was a wide diversity of participants' nationality. The study sample included more than 45 nationalities. Only three groups were comprising above 10 percent of the sample, and these groups were Saudi (nearly 23 percent), British (12 percent), and Chinese (12 percent). The results and a discussion of the different variables and their influence the perceived reading comprehension in terms of demographics are illustrated below.

### **6.3.5 The Impact of Demographic Variables on Reading Comprehension**

#### **6.3.4.1 Effects of gender on reading comprehension**

An Independent Samples t-test was applied to consider the gender effect on the study results. The results showed that there was a strong significant effect on reading comprehension results, with male responders (Mean=12.4828) showing lower performance on Cloze tests than their female counterparts (Mean=13.5959). Significance was found at  $t=-2.289$ ,  $df=364$ ,  $p=0.023$ . As social scientists have generally agreed that, if the probability value, symbolised by the lowercase p, is less than five percent ( $p\text{-value} < 0.05$ ), the result is regarded statistically significant (Nardi, 2015).

**Question One:** Is there a significant relationship between reading comprehension and gender?

The evidence indicated that there was a significant relationship between reading comprehension and gender.

#### **6.3.4.2 Effects of language competence on reading comprehension**

An Independent Samples t-test was applied to investigate the language competence effect on the study research results. The findings indicated a significant effect on reading comprehension results, with English language participants ( $M=15.8646$ ) showing better performance on Cloze tests than their non-native English counterparts ( $M=12.1167$ ). Significance was found at  $t\text{-test}=7.8535$ ,  $df=205$ ,  $p<0.001$ .

**Question Two:** Is there a significant relationship between reading comprehension and language competence

The result displays that the probability value is less than .05, which signifies the result is significant. Therefore, there is a relationship between reading comprehension and language competence.

#### **6.3.4.3 Effects of number of years for studying the English language on reading comprehension**

A one-way ANOVA test - 'f' test- was conducted and it was statistically confirmed that there was a significant effect between study participants based on the number of years of studying English ( $f=6.734$  with significance at  $p<0.001$ ). Tukey's HSD (honest significant difference) test was utilised to identify the source of such significance, and confirmed that there was a significant relationship between the 4 variables; less than 1 year ( $M= 5.8$ ), 1-3 years ( $M= 8.7917$ ), 3-5 years ( $M=11.6825$ ), and more than 7 years ( $M=13.6846$ ), for the favour of the last variable, more than 7 years. Such outcome showed that people with more years of studying the English language were more comprehension of ISPs documents than others with fewer years of studying the English language.

**Question Three:** Is there a significant relationship between reading comprehension and a number of years for studying the English language?

The evidence indicated that there was a significant relationship between reading comprehension and a number of years for studying the English language.

#### **6.3.4.4 Effects of age on reading comprehension**

A one-way ANOVA has been performed and statistically confirmed that there was a significant effect of age on individuals' performance in understanding ISPs texts ( $f=7.039$ ,  $df=3$ ,  $p<0.001$ ). Tukey's HSD test was used to recognise the source of such significance and showed that there was a significant relationship between the two groups, 17-25 and 26-34 ( $p<0.001$ , 95 % Confidence Interval (CI) 1.0134, 3.6582).

**Question Four:** Is there a significant relationship between reading comprehension and age?

According to the one-way ANOVA test finding, there was a significant relationship between reading comprehension and age.

#### **6.3.4.5 Effects of academic qualification on reading comprehension**

A one-way ANOVA test was utilised; for 396 responders as other options were excluded, to explore the significance of the variables. The significant value between the variables was less than 0.01 which indicated that the value had a strong significance and the  $f$ -test =7.258. Tukey's honest significance test was applied to distinguish the significant component(s). Then, the results outlined that there was a significant between all components; Secondary School (M=15.0128), Undergraduate Degree (M=13.0), Master's Degree (M=12.2761), and Doctoral Degree (M=10.7273), for the favour of the first component, Secondary School.

The surprising results indicate an inverse relationship between education level and understanding ISP contents, which was in contrast to some of the readability formulae principals in terms of the ability to comprehend a text based on a reader level of education. Contrariwise, our findings correspond to the results reported by Grant (2010). He expressed that respondents with higher education level were less security aware than other participants with lower education level. Therefore, further statistical analyses were performed in order to know the mystery of the unanticipated results.

One-way ANOVA test results yielded significant value between the variables ( $p=.002$ ) among respective non-native speaker participants only. However, the second statistical analysis was performed for native speaker data and that one-way ANOVA results indicated insignificant value among the study samples' responders. Such results showed that there were significant effects of academic qualification on reading comprehension by only non-native speaker responders.

**Question Five:** Is there a significant relationship between reading comprehension and academic qualification?

The evidence statistically confirmed that there was a significant relationship between reading comprehension and academic qualification.

#### 6.3.4.6 Effects of computer experience level on reading comprehension

A one-way ANOVA test -'f' test- was adopted and it was statistically confirmed that there was a significant effect between study samples' participants based on computer experience level in which  $f=9.088$  with significance at  $p<0.001$ . Tukey method was used to identify the source of such significance, and we found that there was a significant between all components; Proficient ( $M=14.1088$ ), Intermediate ( $M=12.7295$ ), and Basic ( $M=10.6905$ ), for the support of the first component, Proficient. Such finding showed that people with more expertise in computing could perform better. These findings echo the work of Alotaibi et al. (2016).

**Question Six:** Is there a significant relationship between reading comprehension and computer experience level?

According to the one-way ANOVA test finding, there was a significant relationship between reading comprehension and computer experience level.

#### 6.3.4.7 Effects of study subject on reading comprehension

A one-way ANOVA test -'f' test- was conducted and it was statistically confirmed that there was no significant effect between the study samples' participants based on educational discipline.

**Question Seven:** Is there a significant relationship between reading comprehension and study subject?

The evidence indicated that there was no relationship, statistically insignificant, between reading comprehension and gender.

#### 6.3.4.8 Effects of nationality on reading comprehension

A one-way ANOVA test -'f' test- was performed and it was statistically confirmed that there was no significant effect between the study sample's responders based on nationality.

**Question Eight:** Is there a significant relationship between reading comprehension and nationality?

The evidence statistically confirmed that there was no relationship, statistically insignificant, between reading comprehension and nationality.

In summary writing, a t-test has been performed to consider whether there was an effect of gender and language competence on reading comprehension. This indicated that both of them had a significant positive association with reading comprehension. Moreover, a one-way ANOVA test has been performed to consider whether there was an effect of years for studying English language, age, academic qualification, computer experience level, study subject, and nationality on reading comprehension. This indicated that studying English language, age, academic qualification and computer experience level all had a significant positive association with reading comprehension, whereas, study subject and nationality had no significant positive association with reading comprehension, the summary of the results are presented in Table 6.7.

<b>Demographics</b>	<b>Significant Result</b>	<b>Not Significant Result</b>
Gender	✓	
Language Competence	✓	
Years of Studying the English Language	✓	
Age	✓	
Academic Qualification	✓	
Computer Experience Level	✓	
Study Subject		✓
Nationality		✓

Table 6. 7: Analysis results on demographic questions

#### **6.4. Summary**

This chapter has discussed and analysed the collected data of the main study in quantitative method. The results of the main study revealed significant results, which had been compared with a number of readability formulae. This research employed the SPSS/PC statistical package for the purpose of producing a number of tables and bar and pie charts about the demographic profile of the study participants. A summary of major findings of demographic variables was presented in the pre-final section. This chapter also presented the outcomes of the comparison between the human results and the machine results.

The key findings of this chapter will be discussed in the following chapter. In addition, conclusions for this thesis will be provided, as well as presenting any limitations of the research and suggested areas for future research.

## **CHAPTER SEVEN: DISCUSSION AND CONCLUSION**

### **7.1. Introduction**

Along with discussion, this chapter provides the key findings of the research. Additionally, this chapter recommends suggestions to improve users' compliance and concludes with possible limitations and thoughts on future directions of the work.

### **7.2. Discussion**

#### **7.1.1 Summary of the Main Study's Findings**

Several reports have revealed that the ambition of setting up an effective information security policy (ISP) will not be achieved unless users are able to easily become familiar with its content and comply with its requirements. Alotaibi et al. (2016) echo the view that non-compliance with ISPs is one of the significant difficulties confronting institutions. A number of factors enable compliance with regulations and rules of security and one of these is the comprehensibility of the ISP itself. Therefore, in their approach to ISP compliance, institutions should include policy comprehensibility as a significant factor. This aspect can be estimated by applying a bespoke readability formula to the text of ISPs.

Currently, there is no ready mechanism for estimating the likely efficacy of such policies across an organisation. One factor that has a plausible impact upon the comprehensibility of policies is their readability. The aim of this study was to investigate the effectiveness of applying readability metrics as an indicator of likely policy comprehensibility.

Determining readability using only readability formula has limited application. Since the formulae have a number of shortcomings such as assuming readers are alike, only able to measure what can be counted, etc. (see Chapter Three, section 3.4. for more details). Therefore, readability formula software might not accurately reflect the understanding of the reading. Hence, this study measured readability level by readability formulae as well as Cloze tests (readers' responses).

The present study has shown that participants' performance in policy D and policy C was somewhat low (an average of slightly less than six correct answers out of ten). This showed that policy D, as represented by the considered extracts, was the hardest policy to understand for the human reader (which matched the second pilot study results). In contrast, policy G was the easiest to comprehend for respondents. Despite this, all of the policies gained more than 5 out of 10 in mean score, which indicates that respondents (N=396) on average answered correctly more than half of the Cloze test items.

The results of this study indicate that the total average value for the female participants, with a mean of 13.5959, was higher than the male participants, with a mean of 12.4828. This indicates that female performance was better than male performance in understanding security policy concepts. An Independent Samples t-test was applied to consider the gender effect in the study results and showed a strong significant effect on reading comprehension with male responders having lower performance on Cloze tests than their female counterparts. Significance was found at  $t=-2.289$ ,  $df=364$ ,  $p=0.023$ .

In terms of participants' language, 96 people (24.2%) were native speakers of English, while 300 participants (75.8%) were non-native English speakers. The findings revealed that native participants obtained better mean score than their non-native English counterparts. This result is to be expected, as people usually understand their first language better than a foreign language. However, since it is not central to the objectives of this research, such a presumption may be put to the test in future studies. An Independent Samples t-test was utilised to investigate the language competence effect on the study results and findings indicate a significant effect on reading comprehension results with English language participants ( $M=15.8646$ ) having better performance on Cloze tests than their non-native English counterparts ( $M=12.1167$ ). Significance was found at  $t\text{-test}=7.8535$ ,  $df=205$ ,  $p<0.001$ .

In terms of the number of years being taught English, the majority of non-native English participants (49.7% of the total sample) had learned English for more than seven years. In addition, there is a direct relationship between the comprehension of ISP texts and the number of years of learning the English language. Evidently, the



more a person spends time in learning a language the more vocabulary they acquire and the more likely they will comprehend ISP content. However, results raise some concern with those respondents with less than a year or 1-3 years in learning English (nearly 20% of non-native participants). These respondents gained lower than average for the total mean score, which indicates that they may have agreed and signed to ISP documents when they only partially understood. A one-way ANOVA test - 'f' test- was conducted and this confirmed a significant effect between participants' total mean score and the number of years of studying the English language ( $f=6.734$  with significance at  $p<0.001$ ). Tukey's HSD (honest significant difference) test was applied to identify the locus of such significance. This indicated that there is significance between participant's total mean score and the English study period of more than 7 years. It can therefore be assumed that the people with more years of studying English have a greater comprehension of ISP documents than those with fewer years of studying the English language.

Respondents were asked to indicate their age and it emerged that participants in the 55 or over age group had the highest mean rank (15 out of 20). Therefore, the present study raises the possibility that a student in this age group will have more knowledge and experience of ISPs than younger students. Secondly, those aged 17- 25 were in second place for mean rank ( $M=14.051$ ) and they were the majority of participants (49.5% of the total responses). This result may reflect the fact that students at this age are learning or have just learned about computing and security concepts. A one-way ANOVA confirmed that there were significant effects of age on individuals' performance in understanding ISPs texts ( $f=7.039$ ,  $df=3$ ,  $p<0.001$ ). Using Tukey's test to recognise the source of such significance, we found significant results between the two groups, 17-25 and 26-34 ( $p<0.001$ , 95 % CI 1.0134, 3.6582).

In terms of surveyed participant's qualifications, the ranks of the scale mean from the top to bottom were as follows, students of a secondary school ( $M=15.0128$ ), undergraduate degree ( $M=13.0$ ), master degree ( $M=12.2761$ ), and then a doctoral degree ( $M=10.7273$ ). The results revealed that the majority of participants (162) have attained at least graduate level being BSc graduates, 134 respondents were educated to MSc level of education and 78 of the surveyed participants had a secondary school; however, only 22 participants – the smallest group – held a PhD level of education. A

one-way ANOVA test was utilised to explore the significance of the variables. The significance value between the variables was less than 0.01, which indicated that the value had a strong significance and the  $f$ -test = 7.258. Tukey's honest significance test was applied to distinguish the significant component(s). The results indicated significance relationship between all components, Secondary School, Master's Degree, and Doctoral Degree. The surprising results indicated an inverse relationship between education level and understanding ISP contents, which was in contrast to some of the readability formulae principals in terms of the ability to comprehend a text based on a reader's level of education. In contrast to earlier findings, however, our findings correspond to results reported by Grant (2010). He indicated that respondents with higher education level were less security aware than participants with lower education level. Accordingly, further statistical analyses were performed in order to shed light on our unexpected results. One-way ANOVA test results yielded significant value ( $p=0.002$ ) among respective non-native speaker participants only. However, a second statistical analysis was performed for native speaker data and the one-way ANOVA indicated insignificant value among the study samples' responders. Such results showed that there were significant effects of academic qualification on reading comprehension only for non-native speaker responders.

In terms of computer experience, the highest mean score (14.1088) was achieved by 'proficient computer experience' participants. Next in success level were those with an intermediate experience in computing (12.7295 mean score). Third in rank came respondents with only a basic knowledge in the computer field (10.6905 mean score). In effect, computer experience level proved to be directly proportional to the level of comprehension of ISP content so the result of the current study suggests that individuals with more expertise in computing could perform better. A study by Alotaibi et al. (2016) reflects our findings. A one-way ANOVA test -'f' test- was adopted and it was statistically confirmed that there was a significant effect between study samples' participants based on computer experience level in which  $f=9.088$  with significance at  $p<0.001$ . Tukey method was used to identify the source of such significance, and we found that there was a significant between all components; Proficient, Intermediate, and Basic, for the support of the first component, Proficient.

Such finding showed that people with more expertise in computing could perform better.

It is interesting to note that our results across the traditional readability measures revealed that none of the selected traditional formulae give identical scores nor match human rating. The differences in rating the same piece of text are not surprising given previous studies (Klare, 2000; Zhou, Jeong and Green, 2017). The aforementioned comparison between the software readability results and human comprehension test results revealed that the majority of texts' results were matched in human ranking and Strathclyde Readability Measure (SRM). In other words, six out of eight policies gained the same rating in both measurements. The SRM tool did not show any significant distance from the human rating of the documents (the SRM application considers as 'significant distance' if there were three levels difference between user rating and SRM rating). Thereby, we confirm correlations between the software readability formula's results and human comprehension test results, and this supports our view that the readability has an influence on understanding ISPs.

All in all, the insight has drawn from the analysis shows that the results of the selected automated readability formulae (traditional and modern index) for the examined ISP texts were considered as fairly difficult to read by general native English readers. It is worth keeping in mind that these ISP documents are intended for both native and non-native English speakers. This result suggests that there is insufficient attention to ease of comprehension in the process of policy design. This finding is rather disappointing; therefore, we recommend immediate corrective actions to enhance the ease of comprehension for ISPs. This may reduce instances where users avoid fully reading the ISPs and may also increase the likelihood of user compliance.

### **7.2.1 An Indication of the Importance of the Findings**

This is the first study, to our knowledge, to examining the effectiveness of applying readability metrics as an indicator of policy comprehensibility through a number of sequential methods with nearly 450 participants involved in the entire study. The key findings of the study are the agreements in the comprehension test results attributable to the difficulty of the examined texts. The main study showed a strong correlation between the SRM index and human comprehension results and supports our view that

readability has an impact upon understanding ISPs. It was observed that there is a direct relationship between the comprehension of ISP texts and the number of years of learning the English language. This showed that people with more years of studying the English language have a greater comprehension of ISP documents than those with fewer years of studying English. In addition, results raised some concern about respondents with less than a year or 1-3 years in learning English, which represented nearly 20% of non-native participants. These respondents gained lower than average for the total mean score, which indicates that they may have agreed and signed ISP documents when they hardly understood. A significant effect was noted for academic qualification on reading comprehension by non-native speakers. Furthermore, computer experience level was directly proportional to the comprehension of ISP content. Thus, the result implied that individuals with greater expertise in computing could perform better. In terms of the SRM category of examined texts, the results deliver a warning sign to organisations that their readability level is relatively high, with two-thirds of the surveyed policies rated as fairly difficult. We speculate that these results are due to the policy designers not taking account of ease of comprehension when writing ISP documents. We thereby recommend immediate corrective actions to enhance the ease of comprehension for ISPs. This may reduce instances where users avoid reading the ISPs and also increase the likelihood of users' compliance with ISPs.

### **7.3. Contribution and Implications**

There are a number of studies in the literature investigating the factors that enable compliance with regulations and rules of security in order to mitigate security incidents. This research work has five theoretical contributions and practical implications, some of the contributions that I have made are being in one or the other and some are in combined area.

#### **7.3.1 Theoretical and Practical Contributions**

Firstly, this thesis contributed to examining the effectiveness of applying readability metrics as an indicator of policy comprehensibility. This was the first research, to our knowledge, to do this as an experimental study via a number of exploratory sequential mixed methods. The methodological approach of this study is one of the distinguishing characteristics reported in the thesis. Our study depended on

quantitative and qualitative data that have been gathered from many respondents. The research method features qualitative data gathered from experts' opinions and focus group discussions. These data enabled the researcher to design comprehension tests to evaluate human ease of reading on selected texts. Later on, the results from this quantitative approach were analysed to be compared with software readability formulae results. The strength of this work is that the experimental study showed correlations between the SRM's results and human comprehension test results, and this supports our view that readability has an influence on understanding ISPs.

Secondly, although many readability researchers utilise text measurement, our work complements these techniques with comprehension tests as a confirmation of text measurement. This may afford superior understanding for the use of a readability standard as a basis for predicting likely comprehension, particularly for written text.

### **7.3.2 Practical Contributions**

Thirdly, the original data in this thesis were collected from 450 participants by a variety of methods. To our knowledge, there are no existing studies that utilise comprehension test data with large number of participants, since comprehension tests require relatively high engagement from participants.

Fourthly, the study results give a warning sign to all sectors not only to higher education institutions and telecom organisations. Institutions should address policy comprehensibility as they strive to achieve ISP compliance. The readability score of policy documents should be improved by taking an immediate corrective action. In part, this can be estimated by applying a bespoke readability formula to the text of ISPs.

Fifthly, provided guidelines that will permit checks on the likely readability of ISPs and thereby, allow for improvements on reading comprehension. Organisations should regularly check their policy notices and make sure they are understandable by all employees regardless of their level of education. For the general people, governments could run awareness campaigns to raise the understanding of ISPs. Exploiting social media (e.g., Facebook, Twitter), and cross-platform messaging apps

(e.g., WhatsApp, Line) could be a faster and an effective way to increase public awareness.

All in all, this study has three theoretical and practical contributions and two practical contributions that cover the evaluation of readability as a factor in ISPs by software perspective and human perspective.

#### **7.4. Limitations and Suggestions for Future Research**

As with any research, this study has a number of limitations that should be noted. First of all, we only analysed the readability of eight ISPs, this was due to time constraints, as extracting a large amount of data takes substantially more time to be addressed by our sequential methods. Future research could overcome this obstacle by using crowdsourcing platforms such as Amazon's Mechanical Turk (MTurk) to recruit participants, as MTurk is quicker to obtain a larger number of respondents than traditional sampling methods. Yet, future researchers should be wary of the limitations of adopting online platforms to recruit individuals, as mentioned by several scientists (Buhrmester, Kwang and Gosling, 2011; Paolacci and Chandler, 2014; Schmidt, 2015; Follmer, Sperling and Suen, 2017). Secondly, the study addresses two different sectors, public and private (academic and telecom organisations). Future research could extend to other sectors (e.g. banking and financial, insurance, health, etc.).

We have only addressed the readability of online ISPs that were written in English because the SRM formula only accepts English. Future research could extend the SRM tool to include other languages. Another limitation is that the policies were collected in April 2015. Updates by institutions after April 2015 will not be reflected in our data. However, organisations rarely make substantial changes once their ISPs are rolled out. Another limitation of our study, which found that people obtain lower readability score when they are holding higher academic qualifications. Future research could focus on identifying the causes that contributed to the low readability level. Ultimately, our study has also found that respondents with less than a year or 1-3 years in learning English, scored lower than the average of the total mean score. Besides, we have found that two-thirds of the surveyed policies are considered fairly difficult, based on the SRM, is that no advice is afforded on how to improve such documents. Future research could offer insights on how to achieve better readability. For example, sitting down

with policy designers, exploring their opinions, and presenting study results to them may promote evaluation of draft policies for ease of reading prior to policy release.

## **7.5. Recommendations**

Many studies indicate that employee attitudes and lack of security awareness are the most notable contributors to security incidents (Furnell, 2007). Institutions frequently look at increasing users' compliance toward ISPs or 'conditions of use'. This research investigated the effectiveness of applying readability metrics as an indicator of policy comprehensibility. Based on the results of this research, the following recommendations can be applied to future research. First, our study results can be an opportunity for additional research toward a framework to improve users' compliance by accounting for the readability alongside other aspects that influence an individuals' compliance. Second, the study results should undergo further investigation especially for participants' age, and qualification as our research found that respondents within the age group of 26-34 and 35-54 gained lower mean rank than the rest of age groups. Furthermore, our findings indicated an inverse relationship between education level and understanding ISP contents, which is in contrast to some of the readability formula principals in term of the ability to comprehend a text based on a reader level of education. Third, further research on the readability could address policies that were produced in multiple language versions, with a view to enabling establishing their equivalence in terms of readability.

## **7.6. Summary**

This chapter discussed the main outcomes of the research work and summarises the main findings. Moreover, the contribution and implication of the thesis and its originality has been illustrated. Subsequently, limitations and areas for future work were covered to propose suggestions with recommendations for forthcoming research.

## REFERENCES

Abawajy, J., Thatcher, K. and Kim, T. (2008) 'Investigation of stakeholders' commitment to information security awareness programs', in. Busan: Ieee, pp. 472–476. doi: 10.1109/ISA.2008.25.

Abu, Z., Fracgp, H., Mmed, P. S., Fracgp, D. M., Abu Hassan, Z., Schattner, P., Mazza, D., Keluarga, K. and Lumpur, K. (2006) 'Doing a pilot study : why is it essential?', *Malaysian Family Physician*, 1(2), pp. 70–73.

Agarwal, G., Singh, S. and Shukla, R. S. (2010) 'Security analysis of graphical passwords over the alphanumeric passwords', *International Journal of Pure and Applied Sciences and Technology*, 1(2), pp. 60–66. Available at: [http://ijopaasat.in/yahoo\\_site\\_admin/assets/docs/Gaurav\\_paper-5.18191634.pdf](http://ijopaasat.in/yahoo_site_admin/assets/docs/Gaurav_paper-5.18191634.pdf).

Al-Awadi, M. and Renaud, K. (2007) 'Success factors in information security implementation in organizations', in *IADIS International Conference e-Society*. Available at: [https://www.researchgate.net/profile/Karen\\_Renaud/publication/266231077\\_SUCCESS\\_FACTORS\\_IN\\_INFORMATION\\_SECURITY\\_IMPLEMENTATION\\_IN\\_ORGANIZATIONS/links/5492b91c0cf225673b3e083b.pdf](https://www.researchgate.net/profile/Karen_Renaud/publication/266231077_SUCCESS_FACTORS_IN_INFORMATION_SECURITY_IMPLEMENTATION_IN_ORGANIZATIONS/links/5492b91c0cf225673b3e083b.pdf).

ALArifi, A., Tootell, H. and Hyland, P. (2012) 'A study of information security awareness and practices in Saudi Arabia', in. Hammamet: IEEE, pp. 6–12. Available at: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6285845> (Accessed: 27 June 2014).

Albert, W., Tullis, T. and Tedesco, D. (2010) 'Beyond the usability lab: conducting large-scale user experience studies', in *Beyond the Usability Lab*. Morgan Kaufmann, pp. 93–106. Available at: <http://www.sciencedirect.com/science/article/pii/B9780123748928000041/pdfft?md5=05b054f8f24366d6d502121e736601ae&pid=3-s2.0-B9780123748928000041-main.pdf> (Accessed: 17 August 2016).

Alderson, J. . (1983) 'The cloze procedure and proficiency in English as a foreign language', in *Oller (ed.)*, pp. 205–17.

Alkhurayyif, Y. (2015) *Description of the commenting process*. Available at: <https://youtu.be/YciV1m5NYFQ>.

Alnather, M. a. (2015) 'Information security culture critical success factors', in *2015 12th International Conference on Information Technology - New Generations*, pp. 731–735. doi: 10.1109/ITNG.2015.124.

Alnather, M. and Nelson, K. (2009) 'A proposed framework for understanding information security culture and practices in the Saudi context', in. Western Australia: Security Research Centre, pp. 6–17. Available at: <http://ro.ecu.edu.au/cgi/viewcontent.cgi?article=1001&context=ism>.

Alotaibi, F., Furnell, S., Stengel, I. and Papadaki, M. (2016) 'A survey of cybersecurity awareness in Saudi Arabia', in *2016 11th International Conference for Internet Technology and Secured Transactions (ICITST)*. IEEE, pp. 154–158. doi: 10.1109/ICITST.2016.7856687.

Alotaibi, M., Furnell, S. and Clarke, N. (2016) 'Information security policies: a review of challenges and influencing factors', in *2016 11th International Conference*



for *Internet Technology and Secured Transactions (ICITST)*. IEEE, pp. 352–358. doi: 10.1109/ICITST.2016.7856729.

Alreck, P. L. and Settle, R. B. (2004) *The survey research, Handbook (3 Ed.)* New York: McGraw-Hill Education. Available at: <https://www.amazon.com/Survey-Research-Handbook-Third/dp/0072945486>.

Alzamil, Z. A. (2012) ‘Information security awareness at Saudi Arabians’ organizations: an information technology employee’s perspective’, *international journal of information security & Privacy*, 6(3), p. 18. doi: 10.4018/jisp.2012070102.

Amabile, T. M. (1998) ‘How to kill creativity’, *Harvard business review*, 76(5), pp. 76–87. Available at: <http://hbr.org/product/how-to-kill-creativity/an/98501-PDF-ENG>.

Ammann, F. and Sowa, A. (2013) ‘Readability as lever for employees’ compliance with information security policies’, *ISACA*, 4, pp. 1–4. Available at: <http://www.isaca.org/Journal/archives/2013/Volume-4/Documents/13v4-Readability-as-Lever.pdf>.

Amoroso, E. G. (1994) *Fundamentals of computer security technology*. PTR Prentice Hall New Jersey.

Anagnostou, N. and Weir, G. (2006) ‘From corpus-based collocation frequencies to readability measure’, pp. 1–14. Available at: <http://strathprints.strath.ac.uk/2381/>.

Andersen, I. T. (2001) ‘Sicherheit in Europa’, *Status Quo, Trends, Perspektiven*. Studie 2001, Andersen, Dusseldorf.

Anderson, J. M. (2003) ‘Why we need a new definition of information security’, *Computers & Security*. Elsevier, 22(4), pp. 308–313. Available at: <https://www.sciencedirect.com/science/article/pii/S0167404803004073>.

Andress, J. (2014) *The basics of information security: understanding the fundamentals of InfoSec in theory and practice*. Syngress.

De Angeli, A., Coventry, L., Johnson, G. and Renaud, K. (2005) ‘Is a picture really worth a thousand words? exploring the feasibility of graphical authentication systems’, *International Journal of Human-Computer Studies*, 63(1–2), pp. 128–152. doi: 10.1016/j.ijhcs.2005.04.020.

Appleyard, J. (2005) *Information classification: a corporate implementation guide*. Fifth, HF Tipton, & M. Krause, *Information Security Management Handbook*. Fifth. CRC Press LLC.

Ary, D., Jacobs, L. C., Irvine, C. K. S. and Walker, D. (2018) *Introduction to research in education*. Cengage Learning.

Awawdeh, S. Al and Tubaishat, A. (2014) ‘An information security awareness program to address common security concerns in IT unit’, in *2014 11th International Conference on Information Technology: New Generations*. Ieee, pp. 273–278. doi: 10.1109/ITNG.2014.67.

Babbie, E. (2016) *The practical social research*. 14th edn. Wadsworth.

Bailin, A. and Grafstein, A. (2016) *Readability: text and context*. Springer. Available at: <https://link.springer.com/book/10.1057%2F9781137388773>.

Bakhshi, T., Papadaki, M. and Furnell, S. (2009) ‘Social engineering: assessing vulnerabilities in practice’, *Information Management & Computer Security*, 17(1), pp.

53–63. Available at:  
<http://www.emeraldinsight.com/doi/full/10.1108/09685220910944768>.

Barman, S. (2001) *Writing information security policies*. 1st edn. SAMS.

Bartlett, J. E., Kotrlik, J. W. and Higgins, C. C. (2001) ‘Organizational research: determining appropriate sample size in survey research’, *Information technology, learning, and performance journal*. Organizational Systems Research Association, 19(1), p. 43.

Beautement, A. and Sasse, A. (2009) ‘The economics of user effort in information security’, *Computer Fraud & Security*. Elsevier Ltd, 2009(10), pp. 8–12. doi: 10.1016/S1361-3723(09)70127-7.

Bell, J. (2014) *Doing your research project: a guide for first-time researchers*. McGraw-Hill Education (UK).

Van Belle, G. (2011) *Statistical rules of thumb*. John Wiley & Sons.

Benjamin, R. G. (2012) ‘Reconstructing readability: recent developments and recommendations in the analysis of text difficulty’, *Educational Psychology Review*. Springer, 24(1), pp. 63–88. Available at: <https://link.springer.com/content/pdf/10.1007%2Fs10648-011-9181-8.pdf>.

Bequai, A. (1998) ‘Employee abuses in cyberspace: management’s legal quagmire’, *Computers & Security*, 17(8), pp. 667–670. doi: 10.1016/S0167-4048(98)80097-7.

Bernard, M. L., Chaparro, B. S., Mills, M. M. and Halcomb, C. G. (2003) ‘Comparing the effects of text size and format on the readability of computer-displayed Times New Roman and Arial text’, *International Journal of Human-Computer Studies*, 59(6), pp. 823–835. doi: [https://doi.org/10.1016/S1071-5819\(03\)00121-6](https://doi.org/10.1016/S1071-5819(03)00121-6).

Besnard, D. and Arief, B. (2004) ‘Computer security impaired by legitimate users’, *Computers & Security*, 23(3), pp. 253–264. doi: 10.1016/j.cose.2003.09.002.

Biddinika, M. K., Lestari, R. P., Indrawan, B., Yoshikawa, K., Tokimatsu, K. and Takahashi, F. (2016) ‘Measuring the readability of Indonesian biomass websites: the ease of understanding biomass energy information on websites in the Indonesian language’, *Renewable and Sustainable Energy Reviews*, 59, pp. 1349–1357. doi: 10.1016/j.rser.2016.01.078.

Blakley, B., McDermott, E. and Geer, D. (2002) ‘Information security is information risk management’, in *Proceedings of the 2001 workshop on New security paradigms*. ACM, pp. 97–104. Available at: [http://ns2.datacontact.dc.hu/~mfelegyhazi/courses/EconSec/readings/03\\_Blakley2001infosec.pdf](http://ns2.datacontact.dc.hu/~mfelegyhazi/courses/EconSec/readings/03_Blakley2001infosec.pdf).

Bormuth, J. R. (1966) ‘Readability: a new approach’, *Reading research quarterly*. JSTOR, pp. 79–132.

Bormuth, J. R. (1967) ‘Comparable cloze and multiple-choice comprehension test scores’, *Journal of Reading*, 10(5), pp. 291–299. Available at: <http://www.jstor.org/stable/40009351>.

Bormuth, J. R. (1971) *Development of standards of readability: toward a rational criterion of passage performance*. Chicago.

Boujettif, M. and Wang, Y. (2010) ‘Constructivist approach to information

security awareness in the middle east', in. Fukuoka: Ieee, pp. 192–199. doi: 10.1109/BWCCA.2010.70.

Bradburn, N. M., Sudman, S. and Wansink, B. (2004) *Asking questions: the definitive guide to questionnaire design*. Jossey-Bass. Available at: <http://eu.wiley.com/WileyCDA/WileyTitle/productCd-0787970883.html>.

Bray, J., Johns, N. and Kilburn, D. (2011) 'An exploratory study into the factors impeding ethical consumption', *Journal of Business Ethics*, 98(4), pp. 597–608. doi: 10.1007/s10551-010-0640-9.

Briggs, A. R. J., Morrison, M. and Coleman, M. (2012) *Research methods in educational leadership and management*. Sage Publications.

Briney, A. (2000) 'Security focused', *Information Security*, pp. 40–68.

Brostoff, S. and Sasse, M. A. (2000) *Are pass faces more usable than passwords? a field trial investigation*. Edited by S. McDonald, Y. Waern, and G. Cockton. London: Springer London. doi: 10.1007/978-1-4471-0515-2.

Brown, A. S., Bracken, E., Zoccoli, S. and Douglas, K. (2004) 'Generating and remembering passwords', *Applied Cognitive Psychology*, 18(6), pp. 641–651. doi: 10.1002/acp.1014.

Bruce, B., Rubin, A. and Starr, K. (1981) 'Why readability formulas fail', *IEEE Transactions on Professional Communication*, PC-24(1), pp. 50–52. doi: 10.1109/TPC.1981.6447826.

Bryman, A. (2015) *Social research methods*. 5 edition. Oxford university press.

Buhrmester, M., Kwang, T. and Gosling, S. D. (2011) 'Amazon's mechanical Turk', *Perspectives on Psychological Science*, 6(1), pp. 3–5. doi: 10.1177/1745691610393980.

Bulgurcu, B., Cavusoglu, H. and Benbasat, I. (2009) 'Roles of information security awareness and perceived fairness in information security policy compliance', *15th Americas Conference on Information Systems 2009, AMCIS 2009*, 5, pp. 3269–3277. Available at: <http://www.scopus.com/inward/record.url?eid=2-s2.0-84870310775&partnerID=40&md5=0def53089653b82ecede06814596119f>.

Bulgurcu, B., Cavusoglu, H. and Benbasat, I. (2010) 'Information security policy compliance: an empirical study of rationality-based beliefs and information security awareness', *MIS Quarterly*. Society for Information Management and The Management Information Systems Research Center, 34(3), pp. 523–548. Available at: <http://dl.acm.org/citation.cfm?id=2017470.2017477> (Accessed: 10 September 2015).

Burd, S., Ullrich, J., Gavas, E., Kochergin, B., Lehman, L. and Memon, N. (2006) 'Development of the higher education network analysis (HENA) intrusion detection and prevention tool', in *1st Annual Symposium on Information Assurance: Intrusion Detection and Prevention*, p. 9.

Campbell, G. and Weir, G. (2006) 'Matching readers to texts with the Strathclyde readability measure', in *ICT in the Analysis, Teaching and Learning of Languages, Preprints of the ICTATLL Workshop*. Glasgow, UK, pp. 49–55. Available at: <https://pure.strath.ac.uk/portal/files/175302/strathprints002382.pdf>.

Canavan, S. (2003) 'An information security policy development guide for large companies', *SANS Institute*.

Cappelli, D., Moore, A., Trzeciak, R. and Shimeall, T. J. (2009) *Common sense guide to prevention and detection of insider threats*, Published by CERT, Software Engineering Institute, Carnegie Mellon University, <http://www.cert.org>. Available at:

<https://pdfs.semanticscholar.org/0a54/b1b543b32e8ce57887c149c2bf92d986b1c2.pdf>.

Carstens, D. S., Mccauley-bell, P. R., Malone, L. C. and DeMara, R. F. (2004) 'Evaluation of the human impact of password authentication practices on information security', *The International Journal of an Emerging Transdiscipline*, 7(9), pp. 67–85. Available at: <http://www.inform.nu/Articles/Vol7/v7p067-085-229.pdf>.

Caruso, J. B. (2003) 'Information technology security: governance, strategy, and practice in higher education', *EDUCASE Center for Applied Research*, September. Available at: <http://www.personal.psu.edu/kxm/ecar-security.pdf>.

Chan, H. and Mubarak, S. (2012) 'Significance of information security awareness in the higher education sector', *International Journal of Computer Applications*, 60(10), pp. 23–31. Available at: <http://research.ijcaonline.org/volume60/number10/pxc3884202.pdf>.

Chapin, D. a. and Akridge, S. (2005) 'How can security be measured?', *Information Systems Control Journal*, 2, pp. 43–47. Available at: <http://m.isaca.org/Journal/Past-Issues/2005/Volume-2/Documents/jpdf052-how-can-security.pdf>.

Chapple, M. (2005) *Four ways to measure security success*, TechTarget.

Charmaz, K. (2014) *Constructing grounded theory*. Sage.

Chau, S. (2017) *Introducing... street view!*, *Googleblog*. Available at: <https://maps.googleblog.com/2007/05/introducing-street-view.html> (Accessed: 25 May 2018).

Cheek, J. (2004) 'At the margins? discourse analysis and qualitative research', *Qualitative health research*. Sage Publications Sage CA: Thousand Oaks, CA, 14(8), pp. 1140–1150.

Chen, Y., Ramamurthy, K. and Wen, K.-W. (2012) 'Organizations' information security policy compliance: stick or carrot approach?', *Journal of Management Information Systems*, 29(3), pp. 157–188. doi: 10.2753/MIS0742-1222290305.

Cherryholmes, C. H. (1992) 'Notes on pragmatism and scientific realism', *Educational researcher*. Sage Publications Sage CA: Thousand Oaks, CA, 21(6), pp. 13–17.

Child, D. (2017) *Spache readability formula word list*. Available at: <https://readable.io/blog/spache-readability-formula-word-list/> (Accessed: 6 November 2017).

Claburn, T. (2018a) 'Happy having Amazon tiptoe into your house? why not the car, then? In-trunk delivery – what could go wrong?', *The Register*. Available at: [https://www.theregister.co.uk/2018/04/24/amazon\\_wants\\_to\\_drop\\_boxes\\_in\\_your\\_car/](https://www.theregister.co.uk/2018/04/24/amazon_wants_to_drop_boxes_in_your_car/).

Claburn, T. (2018b) 'World's favorite airline' favorite among hackers: British Airways site, app hacked for two weeks, *The Register*. Available at:

[https://www.theregister.co.uk/2018/09/06/british\\_airways\\_hacked/](https://www.theregister.co.uk/2018/09/06/british_airways_hacked/) (Accessed: 9 September 2018).

Clandinin, D. J. and Connelly, F. M. (2000) 'Narrative inquiry: experience and story in qualitative research'. Jossey-Bass San Francisco, CA.

Cohen, L., Manion, L. and Morrison, K. (2011) *Research methods in education*. Routledge.

Collins-Thompson, K. and Callan, J. P. (2004) 'A language modeling approach to predicting reading difficulty', in *Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics: HLT-NAACL 2004*.

Collis, J. and Hussey, R. (2013) *Business research: a practical guide for undergraduate and postgraduate students*. Palgrave macmillan.

Colmer, R. (2017) *The Coleman-Liau index*, [readable.io.com](http://readable.io.com). Available at: <https://readable.io/content/the-coleman-liau-index/> (Accessed: 5 November 2017).

Connelly, L. M. (2008) 'Pilot studies', *MedSurg Nursing*. Jannetti Publications, Inc., 17(6), pp. 411–413.

Corbin, J. and Strauss, A. (2015) *Basics of qualitative research: techniques and procedures for developing grounded theory*. 4 edition. Sage Publication.

Couper, M. P. (2000) 'Review web surveys: a review of issues and approaches', *The Public Opinion Quarterly*. JSTOR, 64(4), pp. 464–494.

Cox, A., Connolly, S. and Currall, J. (2001) 'Raising information security awareness in the academic setting', *Vine*. MCB UP Ltd, 31(2), pp. 11–16. Available at: <http://www.emeraldinsight.com/journals.htm?issn=0305-5728&volume=31&issue=2&articleid=1502425&show=pdf&PHPSESSID=8t5m78pf6tgV6hq9g8o8ss1fq1>.

Craig, R. (2011) *Security awareness and training programme design: a case study*. Available at: <http://www.computerweekly.com/tip/Security-awareness-and-training-programme-design-A-case-study> (Accessed: 30 September 2014).

Creswell, J. W. (2012) *Educational research: planning, conducting, and evaluating quantitative and qualitative research*. 4 edition. Pearson. Available at: <http://basu.nahad.ir/uploads/creswell.pdf>.

Creswell, J. W. (2013) *Research design (international student edition): qualitative, quantitative, and mixed methods approaches*. 4 edition. SAGE Publications.

Creswell, J. W. (2015) *A concise introduction to mixed methods research*. Sage Publications.

Creswell, J. W. and Clark, V. L. P. (2007) 'Designing and conducting mixed methods research'. Wiley Online Library.

Creswell, J. W. and Creswell, J. D. (2018) *Research design: qualitative, quantitative, and mixed methods approaches*. 5 edition. AGE Publications Inc.

Creswell, J. W. and Poth, C. N. (2017) *Qualitative inquiry and research design: choosing among five approaches*. 4 edition. SAGE Publications.

Crosby, D. (2016) *Methodology for eliciting expert opinion*, [Mrc.ac.uk](http://mrc.ac.uk). Available at: <https://www.mrc.ac.uk/funding/how-we-fund-research/highlight->

notices/methodology-for-eliciting-expert-opinion/ (Accessed: 3 March 2016).

Crossley, S. A., Dufty, D. F., McCarthy, P. M. and McNamara, D. S. (2007) 'Toward a new readability: a mixed model approach', in *Proceedings of the Annual Meeting of the Cognitive Science Society*, pp. 197–202. Available at: <https://cloudfront.escholarship.org/dist/prd/content/qt39r3d755/qt39r3d755.pdf>.

Crossley, S. A., Skalicky, S., Dascalu, M., McNamara, D. S. and Kyle, K. (2017) 'Predicting text comprehension, processing, and familiarity in adult readers: new approaches to readability formulas', *Discourse Processes*. Routledge, 00(00), pp. 1–20. doi: 10.1080/0163853X.2017.1296264.

D'Arcy, J. and Greene, G. (2014) 'Security culture and the employment relationship as drivers of employees' security compliance', *Information Management & Computer Security*, 22(5), pp. 474–489. doi: 10.1108/IMCS-08-2013-0057.

Darroch, I., Goodman, J., Brewster, S. and Gray, P. (2005) 'The effect of age and font size on reading text on handheld computers', in Costabile, M. F. and Paternò, F. (eds) *Human-Computer Interaction - INTERACT 2005*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 253–266. Available at: [https://link.springer.com/chapter/10.1007/11555261\\_23](https://link.springer.com/chapter/10.1007/11555261_23).

David, J. (2002) 'Policy enforcement in the workplace', *Computers & Security*, 21(6), pp. 506–513. doi: [https://doi.org/10.1016/S0167-4048\(02\)01006-4](https://doi.org/10.1016/S0167-4048(02)01006-4).

David, M. and Sutton, C. D. (2004) *Social research: the basics*. Sage.

Denning, D. E. R. (1999) *Information warfare and security*. Addison-Wesley Reading, MA. Available at: [http://nnt.es/Information warfare and security.pdf](http://nnt.es/Information%20warfare%20and%20security.pdf).

Denzin, N. K. and Lincoln, Y. S. (2018) *The Sage handbook of qualitative research*. 5 edition. Sage.

Dhillon, G. (2001) 'Challenges in managing information security in the new millennium', in *Information security management: Global challenges in the new millennium*. IGI Global, pp. 1–8. Available at: <https://www.igi-global.com/chapter/information-security-management/23356>.

Dhillon, G. and Backhouse, J. (1996) 'Risks in the use of information technology within organizations', *International Journal of Information Management*. Pergamon, 16(1), pp. 65–74. Available at: <https://pdfs.semanticscholar.org/eb3f/bce4e5e630cd114d6a9fd54e1c14318b2904.pdf>.

Dhillon, G. and Backhouse, J. (2000a) 'Information system security management in the new millennium', *Communications of the ACM*. ACM, 43(7), pp. 125–128. Available at: <http://www.academia.edu/download/26008100/luento1.pdf>.

Dhillon, G. and Backhouse, J. (2000b) 'Technical opinion: information system security management in the new millennium', *Communications of the ACM*. ACM, 43(7), pp. 125–128. Available at: [https://s3.amazonaws.com/academia.edu.documents/26008100/luento1.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1537106649&Signature=tZpLYLAR78gg4xcqnpfxI4pwSo4%3D&response-content-disposition=inline%3Bfilename%3DTechnical\\_opinion\\_Information\\_system\\_sec](https://s3.amazonaws.com/academia.edu.documents/26008100/luento1.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1537106649&Signature=tZpLYLAR78gg4xcqnpfxI4pwSo4%3D&response-content-disposition=inline%3Bfilename%3DTechnical_opinion_Information_system_sec).

Dillman, D. A., Smyth, J. D. and Christian, L. M. (2014) *Internet, phone, mail, and*

*mixed-mode surveys: the tailored design method*. John Wiley & Sons.

Doherty, N. F. and Fulford, H. (2005) 'Do information security policies reduce the incidence of security breaches: an exploratory analysis', *Information Resources Management Journal (IRMJ)*. IGI Global, 18(4), pp. 21–39. Available at: [https://www.researchgate.net/profile/Neil\\_Doherty2/publication/280697855\\_Information\\_Security\\_Policy\\_Effectiveness\\_-\\_repository\\_copy\\_2011/data/55c1b39408aec0e5f4491e19/Information-Security-Policy-Effectiveness-repository-copy-2011.doc](https://www.researchgate.net/profile/Neil_Doherty2/publication/280697855_Information_Security_Policy_Effectiveness_-_repository_copy_2011/data/55c1b39408aec0e5f4491e19/Information-Security-Policy-Effectiveness-repository-copy-2011.doc).

Dourish, P., Grinter, E., Delgado De La Flor, J. and Joseph, M. (2004) 'Security in the wild: user strategies for managing security as an everyday, practical problem', *Personal and Ubiquitous Computing*. Springer-Verlag, 8(6), pp. 391–401. Available at: <https://link.springer.com/content/pdf/10.1007/s00779-004-0308-5.pdf>.

DuBay, W. (2004) *The principles of readability*, Costa Mesa: Impact Information. doi: 10.1.1.91.4042.

Edwards, R. P. A. and Gibbon, V. (1973) *Words your children use*. London: Burke Books.

Ellering, N. (2016) *What 10 studies say about the best time to send email, coschedule*. Available at: <https://coschedule.com/blog/best-time-to-send-email/> (Accessed: 22 April 2018).

ENISA (2007) 'Current practice and the measurement of success', *European Network and Information Security Agency*, (July), p. 20. Available at: [https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0CCMQFjAA&url=http://www.eicar.org/files/enisa\\_measuring\\_awareness.pdf&ei=bNduVZynBabnygOUtoO4Dw&usg=AFQjCNFccXyg8phCf7QTWE5RPDGqpXQziQ&sig2=A6eJvXMLepnxPK0fy\\_IDYQ&bvm=bv.9](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0CCMQFjAA&url=http://www.eicar.org/files/enisa_measuring_awareness.pdf&ei=bNduVZynBabnygOUtoO4Dw&usg=AFQjCNFccXyg8phCf7QTWE5RPDGqpXQziQ&sig2=A6eJvXMLepnxPK0fy_IDYQ&bvm=bv.9).

ENISA (2010) *The new users' guide: how to raise information security awareness*. ENISA. doi: 10.2824/19110.

Ernest Chang, S. and Lin, C.-S. (2007) 'Exploring organizational culture for information security management', *Industrial Management & Data Systems*. Emerald Group Publishing Limited, 107(3), pp. 438–458. Available at: [https://www.researchgate.net/profile/Shuchih\\_Ernest\\_Chang/publication/220672511\\_Exploring\\_organizational\\_culture\\_for\\_information\\_security\\_management/links/5630edba08ae13bc6c3549a3/Exploring-organizational-culture-for-information-security-management.pdf](https://www.researchgate.net/profile/Shuchih_Ernest_Chang/publication/220672511_Exploring_organizational_culture_for_information_security_management/links/5630edba08ae13bc6c3549a3/Exploring-organizational-culture-for-information-security-management.pdf).

Ernst and Young (2012) *Fighting to close the gap Global information security survey 2012*. Available at: [http://www.ey.com/Publication/vwLUAssets/GISS2012/\\$FILE/EY\\_GISS\\_2012.pdf](http://www.ey.com/Publication/vwLUAssets/GISS2012/$FILE/EY_GISS_2012.pdf).

Ernst and Young (2018) *Cybersecurity regained: preparing to face cyber attacks 20th global information security survey 2017-18*. Available at: [https://www.ey.com/Publication/vwLUAssets/ey-cybersecurity-regained-preparing-to-face-cyber-attacks/\\$FILE/ey-cybersecurity-regained-preparing-to-face-cyber-attacks.pdf](https://www.ey.com/Publication/vwLUAssets/ey-cybersecurity-regained-preparing-to-face-cyber-attacks/$FILE/ey-cybersecurity-regained-preparing-to-face-cyber-attacks.pdf).

Evaluation Research Team (2008) 'Data collection methods for program evaluation: focus groups', *Evaluation Briefs*, (13), p. 2. Available at: <http://www.cdc.gov/healthyyouth/evaluation/index.htm>.

Evening Standard (2007) 'Nationwide fined £1m over stolen laptop scandal', *Evening Standard*. Available at: <https://www.standard.co.uk/news/nationwide-fined-1m-over-stolen-laptop-scandal-7249203.html>.

Feng, L., Elhadad, N. and Huenerfauth, M. (2009) 'Cognitively motivated features for readability assessment', in *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics*. Stroudsburg, PA, USA: Association for Computational Linguistics (EACL '09), pp. 229–237. Available at: <http://dl.acm.org/citation.cfm?id=1609067.1609092>.

Fetterman, D. M. (2010) *Ethnography: step-by-step*. Sage.

Fine, E. and Handelsman, J. (2010) 'Benefits and challenges of diversity in academic settings', *the Board of Regents of the University of Wisconsin System*. Available at: [https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwioiKvQmvLKAhWGwBQKHV9hCQwQFgghMAA&url=http://wiseli.engr.wisc.edu/docs/Benefits\\_Challenges.pdf&usg=AFQjCNF28PmqtNZKZlQvd8PNMKojEv8Guw&sig2=Y9k1CKNxQk1aBneQjmlsdw&bvm](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwioiKvQmvLKAhWGwBQKHV9hCQwQFgghMAA&url=http://wiseli.engr.wisc.edu/docs/Benefits_Challenges.pdf&usg=AFQjCNF28PmqtNZKZlQvd8PNMKojEv8Guw&sig2=Y9k1CKNxQk1aBneQjmlsdw&bvm).

Fitzgerald, T. (2016) *Information security governance simplified: from the boardroom to the keyboard*. CRC Press.

Follmer, D. J., Sperling, R. A. and Suen, H. K. (2017) 'The role of MTurk in education research: advantages, issues, and future directions', *Educational Researcher*. American Educational Research Association, p. 0013189X17725519. doi: 10.3102/0013189X17725519.

Foltz, P. W., Kintsch, W. and Landauer, T. K. (1998) 'The measurement of textual coherence with latent semantic analysis', *Discourse processes*. Taylor & Francis, 25(2–3), pp. 285–307. Available at: <http://www.tandfonline.com/doi/abs/10.1080/01638539809545029?needAccess=true#aHR0cDovL3d3dy50YW5kZm9ubGluZS5jb20vZG9pL3BkZi8xMC4xMDgwLzAxNjM4NTM5ODA5NTQ1MDI5P25lZWRY2Nlc3M9dHJlZUBAQAQDA=>.

Fuchs, L. S., Fuchs, D. and Maxwell, L. (1988) 'The validity of informal reading comprehension measures', *Remedial and Special Education*, pp. 20–28. doi: 10.1177/074193258800900206.

Fulford, H. and Doherty, N. F. (2003) 'The application of information security policies in large UK-based organizations: an exploratory investigation', *Information Management & Computer Security*. MCB UP Ltd, 11(3), pp. 106–114. doi: 10.1108/09685220310480381.

Furnell, S. (2007) 'IFIP workshop – information security culture', *Computers & Security*, 26(1), p. 35. doi: 10.1016/j.cose.2006.10.012.

Furnell, S. and Clarke, N. (2012) 'Power to the people? the evolving recognition of human aspects of security', *Computers & Security*. Elsevier, 31(8), pp. 983–988.

Garbars, K. (2002) 'Implementing an effective IT security program', *SANS Institute*. Retrieved January, 11, p. 2004. Available at: <https://www.sans.org/reading-room/whitepapers/bestprac/implementing-effective-security-program-80>.

Gartenberg, C. (2018) 'Google Maps is getting augmented reality directions and recommendation features', *Theverge*. Available at: <https://www.theverge.com/2018/5/8/17332480/google-maps-augmented-reality->



directions-walking-ar-street-view-personalized-recommendations-voting.

Gaunt, N. (1998) 'Installing an appropriate information security policy', *International Journal of Medical Informatics*. Elsevier, 49(1), pp. 131–134.

Georgetown University (2015) *Top 10 threats to information security*, Georgetown University. Available at: <https://sconline.georgetown.edu/programs/masters-technology-management/resources/top-threats-to-information-technology> (Accessed: 13 August 2018).

Glassner, B. and Moreno, J. D. (2013) *The qualitative-quantitative distinction in the social sciences*. Springer Science & Business Media.

Gold, K. (2014) *No more excuses: encrypt your laptops or pay big \$*, Mintz Levin. Available at: <https://www.healthlawpolicymatters.com/2014/04/24/no-more-excuses-encrypt-your-laptops-or-pay-big/> (Accessed: 14 August 2018).

Gollmann, D. (2010) 'Computer security', *Wiley Interdisciplinary Reviews: Computational Statistics*. Wiley Online Library, 2(5), pp. 544–554. Available at: <https://onlinelibrary.wiley.com/doi/epdf/10.1002/wics.106>.

Gonzalez, J. J. and Sawicka, A. (2002) 'A framework for human factors in information security', in *Wseas international conference on information security, Rio de Janeiro*, pp. 187–448. Available at: [https://www.researchgate.net/profile/Jose\\_Gonzalez74/publication/228684288\\_A\\_framework\\_for\\_human\\_factors\\_in\\_information\\_security/links/02e7e53720ad4cfbb3000000.pdf](https://www.researchgate.net/profile/Jose_Gonzalez74/publication/228684288_A_framework_for_human_factors_in_information_security/links/02e7e53720ad4cfbb3000000.pdf).

Gorman, G. E., Clayton, P. R., Shep, S. J. and Clayton, A. (2005) *Qualitative research for the information professional: a practical handbook*. Facet Publishing.

Graesser, A. C., McNamara, D. S., Cai, Z. and McCarthy, P. M. (2014a) 'Coh-Metrix measures', in Graesser, A. C., McNamara, D. S., McCarthy, P. M., and Cai, Z. (eds) *Automated Evaluation of Text and Discourse with Coh-Metrix*. Cambridge: Cambridge University Press, pp. 60–77. doi: DOI: 10.1017/CBO9780511894664.006.

Graesser, A. C., McNamara, D. S., Cai, Z. and McCarthy, P. M. (2014b) 'Coh-Metrix measures of text readability and easability', in Graesser, A. C., McNamara, D. S., McCarthy, P. M., and Cai, Z. (eds) *Automated Evaluation of Text and Discourse with Coh-Metrix*. Cambridge: Cambridge University Press, pp. 78–95. doi: DOI: 10.1017/CBO9780511894664.007.

Graesser, A. C., McNamara, D. S., Louwerse, M. M. and Cai, Z. (2004) 'Coh-Metrix: analysis of text on cohesion and language', *Behavior research methods, instruments, & computers*. Springer, 36(2), pp. 193–202. Available at: <https://link.springer.com/content/pdf/10.3758%2FBF03195564.pdf>.

Grant, G. J. (2010) *Ascertaining the relationship between security awareness and the security behavior of individuals*. Available at: <http://search.proquest.com/pqdtft/docview/759229194/fulltextPDF/1CBE1B4B68F54137PQ/1?accountid=142908>.

Gray, D. E. (2013) *Doing research in the real world*. Sage.

Gray, W. S. and Leary, B. (1935) *What makes a book readable*. Chicago: University of Chicago Press. Available at: <http://archive.org/stream/whatmakesabookre028092mbp/whatmakesabookre028092>

mbp\_djvu.txt.

Greenwald, S. J. (1999) 'Discussion topic: what is the old security paradigm?', in *Proceedings of the 1998 workshop on New security paradigms*. ACM, pp. 107–118. Available at: [https://dl.acm.org/ft\\_gateway.cfm?id=310925&ftid=18459&dwn=1&CFID=46150278&CFTOKEN=a173d9d440fbe597-6289C27D-EFC8-3103-0F5550F9EE8A638B](https://dl.acm.org/ft_gateway.cfm?id=310925&ftid=18459&dwn=1&CFID=46150278&CFTOKEN=a173d9d440fbe597-6289C27D-EFC8-3103-0F5550F9EE8A638B).

Gritzalis, D. (1997) 'A baseline security policy for distributed healthcare information systems', *Computers & Security*. Elsevier, 16(8), pp. 709–719. Available at: [https://scholar.google.co.uk/scholar?output=instlink&q=info:wqoPNXImGN0J:scholar.google.com/&hl=en&as\\_sdt=0,5&scillfp=13062347873467713956&oi=lle](https://scholar.google.co.uk/scholar?output=instlink&q=info:wqoPNXImGN0J:scholar.google.com/&hl=en&as_sdt=0,5&scillfp=13062347873467713956&oi=lle).

Gross, J. B. and Rosson, M. B. (2007) 'Looking for trouble', in *Proceedings of the 2007 symposium on Computer human interaction for the management of information technology - CHIMIT '07*. New York, New York, USA: ACM Press, p. 10. doi: 10.1145/1234772.1234786.

Guba, E. G. (1990) *The paradigm dialog*. Sage publications.

Guillemette, R. A. (1989) 'The Cloze procedure: assessing the understandability of an IEEE standard', *IEEE Transactions on Professional Communication*, 32(1), pp. 41–47. doi: 10.1109/47.21861.

Gupta, A. and Hammond, R. (2005) 'Information systems security issues and decisions for small businesses: an empirical examination', *Information management & computer security*. Emerald Group Publishing Limited, 13(4), pp. 297–310.

Hanley, M., Dean, T., Schroeder, W., Houy, M., Trzeciak, R. F. and Montelibano, J. (2011) *An analysis of technical observations in insider theft of intellectual property cases*, *Software Engineering Institute*. Available at: <http://repository.cmu.edu/sei/666/> (Accessed: 10 July 2017).

Harris, M. M. and Schaubroeck, J. (1990) 'Confirmatory modelling in organizational behavior/human resource management: issues and applications', *Journal of Management*. SAGE Publications, 16(2), pp. 337–360. doi: 10.1177/014920639001600206.

Hentea, M., Dhillon, H. S. and Dhillon, M. (2006) 'Towards changes in information security education', *Journal of Information Technology Education: Research*. Informing Science Institute, 5, pp. 221–233. Available at: [https://www.learntechlib.org/p/111542/article\\_111542.pdf](https://www.learntechlib.org/p/111542/article_111542.pdf).

Herath, T. and Rao, H. R. (2009) 'Protection motivation and deterrence: a framework for security policy compliance in organisations', *European Journal of Information Systems*, 18(2), pp. 106–125. doi: 10.1057/ejis.2009.6.

HESA (2017) *Introduction - students 2015/16*, *Higher Education Statistics Agency*. Available at: <https://www.hesa.ac.uk/data-and-analysis/publications/students-2015-16/introduction> (Accessed: 18 April 2018).

HESA (2018) *Higher education student statistics: UK 2016/17 - student numbers and characteristics*, *Higher Education Statistics Agency*. Available at: <https://www.hesa.ac.uk/news/11-01-2018/sfr247-higher-education-student-statistics/numbers> (Accessed: 18 April 2018).

Higgins, C. A. and Compeau, D. R. (1995) 'Development of a measure and initial test', *MIS Quarterly*, 19(2), pp. 189–211. doi: 10.2307/249688.

Higgins, N. H. (1999) 'Corporate system security: towards an integrated management approach', *Information Management & Computer Security*, 7(5), pp. 217–222. doi: 10.1108/09685229910292817.

Hill, R. (1998) 'What sample size is "enough" in internet survey research', *Interpersonal Computing and Technology: An electronic journal for the 21st century*, 6(3–4), pp. 1–12. Available at: <http://reconstrue.co.nz/IPCT-J Vol 6 Robin hill SampleSize.pdf>.

Hill, R. (2017) 'Creepy Cayla doll violates liberté publique, screams French data protection agency', *The Register*. Available at: [https://www.theregister.co.uk/2017/12/04/creepy\\_cayla\\_doll\\_breaches\\_french\\_data\\_protection\\_rules\\_says\\_agency/](https://www.theregister.co.uk/2017/12/04/creepy_cayla_doll_breaches_french_data_protection_rules_says_agency/).

Hinde, S. (2002) 'Security surveys spring crop', *Computers & Security*. Elsevier, 21(4), pp. 310–321.

Holbert, D. A. (2013) *Factors contributing to security awareness of the end user*. Capella University. Available at: <http://search.proquest.com/pqdft/docview/1476209084/fulltextPDF/2065219B6B5F4E63PQ/1?accountid=142908>.

Höne, K. and Eloff, J. H. . (2002a) 'Information security policy—what do international information security standards say?', *Computers & Security*. Elsevier, 21(5), pp. 402–409.

Höne, K. and Eloff, J. H. . (2002b) 'What makes an effective information security policy?', *Network Security*, 2002(6), pp. 14–16. doi: 10.1016/S1353-4858(02)06011-7.

Hong, K., Chi, Y., Chao, L. R. and Tang, J. (2006) 'An empirical study of information security policy on information security elevation in Taiwan', *Information Management & Computer Security*. Emerald, 14(2), pp. 104–115. doi: 10.1108/09685220610655861.

Huff, S. L. and Munro, M. C. (1985) 'Information technology assessment and adoption: a field study', *MIS quarterly*. JSTOR, pp. 327–340. Available at: <https://www.jstor.org/stable/pdf/249233.pdf?refreqid=excelsior%3A969da610b772afcdc8d4ae5eee6f59e3>.

*InfoSec skills* (2018). Available at: <https://academy.infosecskills.com//login/> (Accessed: 29 August 2018).

Isaac, S. and Michael, W. (1995) *Handbook in research and evaluation*. San Diego: CA: Educational and Industrial Testing Services.

ISACA (2009) 'An introduction to the business model for information security', *Information Security*, pp. 1–28. Available at: [http://www.isaca.org/Knowledge-Center/Research/Documents/Introduction-to-the-Business-Model-for-Information-Security\\_res\\_Eng\\_0109.pdf](http://www.isaca.org/Knowledge-Center/Research/Documents/Introduction-to-the-Business-Model-for-Information-Security_res_Eng_0109.pdf).

ISO/IEC (2018) *Information technology — security techniques — information security management systems — overview and vocabulary*. Available at: <https://www.iso.org/obp/ui/fr/#iso:std:iso-iec:27000:ed-5:v1:en>.

ISO (2001) *Information Technology: code of practice for information security management*. London: British Standards Institution.

Jackson, K. M., Pukys, S., Castro, A., Hermosura, L., Mendez, J., Vohra-Gupta, S., Padilla, Y. and Morales, G. (2018) 'Using the transformative paradigm to conduct a mixed methods needs assessment of a marginalized community: methodological lessons and implications', *Evaluation and Program Planning*. Pergamon, 66, pp. 111–119. doi: <https://doi.org/10.1016/j.evalprogplan.2017.09.010>.

Janan, D. and David, W. (2012) 'Readability: the limitations of an approach through formulae', *British Educational Research Association Annual Conference, University of Manchester*, pp. 1–16. Available at: <http://www.leeds.ac.uk/educol/documents/213296.pdf>.

Janssens, W., De Pelsmacker, P. and Van Kenhove, P. (2008) *Marketing research with SPSS*. Pearson Education.

Jayaratra, N. (1994) *Understanding and evaluating methodologies: NIMSAD, a systematic framework*. McGraw-Hill, Inc.

Jayaratra, Y. S. N., Anderson, N. K. and Zwahlen, R. A. (2014) 'Readability of websites containing information on dental implants', *Clinical oral implants research*. Wiley Online Library, 25(12), pp. 1319–1324. Available at: [https://www.researchgate.net/profile/Roger\\_Zwahlen/publication/258034432\\_Readability\\_of\\_websites\\_containing\\_information\\_on\\_dental\\_implants/links/5739131208aea45ee83ec2d6.pdf](https://www.researchgate.net/profile/Roger_Zwahlen/publication/258034432_Readability_of_websites_containing_information_on_dental_implants/links/5739131208aea45ee83ec2d6.pdf).

Johanson, G. A. and Brooks, G. P. (2010) 'Initial scale development: sample size for pilot studies', *Educational and Psychological Measurement*, 70(3), pp. 394–400. doi: 10.1177/0013164409355692.

Johnson, B. and Turner, L. A. (2003) 'Data collection strategies in mixed methods research', *Handbook of mixed methods in social and behavioral research*, pp. 297–319.

Johnson, E. C. (2006) 'Security awareness: switch to a better programme', *Network Security*, pp. 15–18. doi: 10.1016/S1353-4858(06)70337-3.

Johnson, R. B. and Onwuegbuzie, A. J. (2004) 'Mixed methods research: a research paradigm whose time has come', *Educational researcher*. Sage Publications Sage CA: Thousand Oaks, CA, 33(7), pp. 14–26.

Jonker, J. and Pennink, B. (2010) *The essence of research methodology: a concise guide for master and PhD students in management science*. Springer Science & Business Media.

Julious, S. A. (2005) 'Sample size of 12 per group rule of thumb for a pilot study', *Pharmaceutical Statistics*, 4(4), pp. 287–291. doi: 10.1002/pst.185.

Kajzer, M., D'Arcy, J., Crowell, C. R., Striegel, A. and Van Bruggen, D. (2014) 'An exploratory investigation of message-person congruence in information security awareness campaigns', *Computers & Security*. Elsevier, 43, pp. 64–76. doi: 10.1016/j.cose.2014.03.003.

Kankanhalli, A., Teo, H.-H., Tan, B. C. Y. and Wei, K.-K. (2003) 'An integrative study of information systems security effectiveness', *International Journal of Information Management*, 23(2), pp. 139–154. doi: 10.1016/S0268-4012(02)00105-6.

Katz, F. H. (2005) 'The effect of a university information security survey on instruction methods in information security', in *Proceedings of the 2nd annual conference on Information security curriculum development*. ACM, pp. 43–48.

Kemmis, S. and McTaggart, R. (2005) *Participatory action research: communicative action and the public sphere*. Sage Publications Ltd.

Khan, B., Alghathbar, K. S., Nabi, S. I. and Khan, M. K. (2011) 'Effectiveness of information security awareness methods based on psychological theories', *African Journal of Business Management*. Academic Journals, 5(26), pp. 10862–10868. Available at: <http://www.academicjournals.org/journal/AJBM/article-full-text-pdf/8CF5EAB16308>.

Kirsch, L. and Boss, S. (2007) 'The last line of defense: motivating employees to follow corporate security guidelines', *ICIS 2007 Proceedings*, p. 103. Available at: <https://pdfs.semanticscholar.org/6822/0f29e24ecac20e36aa99c6dc32b7d4a4f6d7.pdf>.

Klare, G. R. (1985) *How to write readable English*. Hutchinson.

Klare, G. R. (2000) 'Readable computer documentation', *ACM J. Comput. Doc.* New York, NY, USA: ACM, 24(3), pp. 148–168. doi: 10.1145/344599.344645.

Klenke, K. (2016) *Qualitative research in the study of leadership*. Emerald Group Publishing Limited.

Kline, R. B. and Santor, D. A. (1999) 'Principles & practice of structural equation modelling', *Canadian Psychology*. Canadian Psychological Association, 40(4), p. 381.

Knapton, S. (2018) 'Hackers could kill patients by attacking their pacemakers, warns Royal Academy of Engineering', *elegraph Media Group*. Available at: <https://www.telegraph.co.uk/science/2018/03/14/hackers-could-kill-patients-attacking-pacemakers-warns-royal/>.

Kobayashi, M. (2009) *Hitting the mark: how can text organisation and response format affect reading test performance?* Peter Lang.

Kobayashi, M. (2012) 'Cloze tests revisited: exploring item characteristics with special attention to scoring methods', *The Modern Language Journal*. Blackwell Publishers Inc., 86(4), pp. 571–586. doi: 10.1111/1540-4781.00162.

Kochetkova, K. (2015) *Allegedly 40 apps on app store are infected, Kaspersky*. Available at: [https://usa.kaspersky.com/blog/xcodeghost-compromises-apps-in-app-store/6030/?\\_ga=2.36941600.1287548340.1534257108-797297430.1534257108](https://usa.kaspersky.com/blog/xcodeghost-compromises-apps-in-app-store/6030/?_ga=2.36941600.1287548340.1534257108-797297430.1534257108) (Accessed: 14 August 2018).

Kohn, A. (1993) 'Why incentive plans cannot work', *Harvard Business Review*, 71(5), pp. 54–63. doi: Article.

Kowalski, E., Cappelli, D. and Moore, A. (2008) *Insider threat study: illicit cyber activity in the information technology and telecommunications sector*. CARNEGIE-MELLON UNIV PITTSBURGH PA SOFTWARE ENGINEERING INST. Available at: <http://www.dtic.mil/get-tr-doc/pdf?AD=ADA638653>.

Krebs, B. (2014) *The target breach, by the numbers, Krebs on Security RSS*. Available at: <https://krebsonsecurity.com/2014/05/the-target-breach-by-the-numbers/> (Accessed: 13 August 2018).

Krejcie, R. V and Morgan, D. W. (1970) 'Determining sample size for research activities', *Educ psychol meas*. ERIC. Available at: <https://eric.ed.gov/?id=EJ026025>.

Krueger, R. A. and Casey, M. A. (2014) *Focus groups: a practical guide for applied research*. SAGE Publications.

Kumar, M., Garfinkel, T., Boneh, D. and Winograd, T. (2007) 'Reducing shoulder-surfing by using gaze-based password entry', in *Proceedings of the 3rd symposium on Usable privacy and security - SOUPS '07*. New York, New York, USA: ACM Press, p. 13. doi: 10.1145/1280680.1280683.

Lampson, B. W. (2004) 'Computer security in the real world', *Computer*. IEEE, 37(6), pp. 37–46. Available at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.66.4301&rep=rep1&type=pdf>.

Lane, T. (2007) 'Information security management in Australian universities: an exploratory analysis'. Queensland University of Technology. Available at: [http://eprints.qut.edu.au/16486/1/Tim\\_Lane\\_Thesis.pdf](http://eprints.qut.edu.au/16486/1/Tim_Lane_Thesis.pdf).

Leach, J. (2003) 'Improving user security behaviour', *Computers & Security*. Elsevier, 22(8), pp. 685–692. Available at: <https://www.sciencedirect.com/science/article/pii/S0167404803000075>.

LearnClick Blog (2016) *LearnClick blog*. Available at: <https://www.learnclick.com/blog/> (Accessed: 30 October 2016).

Lebek, B., Uffen, J., Breitner, M. H., Neumann, M. and Hohler, B. (2013) 'Employees' information security awareness and behavior: a literature review', in. Wailea, HI, USA: Ieee, pp. 2978–2987. doi: 10.1109/HICSS.2013.192.

De Leeuw, E., J. Hox, J. and Dillman, D. A. (2012) *International handbook of survey methodology*. New York, USA: Routledge Ltd - M.U.A. Available at: <https://www.dawsonera.com/abstract/9780203843123>.

Leicester University (2016) *How much do you really know?* Available at: <https://eu1.gomolearning.com/builds/20825/preview/5819ba3a24e67/index.htm?gomPreview=true&> (Accessed: 29 August 2018).

Leyden, J. (2015) 'Crumbs! stricken kiev blames Russian hackers for Xmas eve outages', *The Register*. Available at: [https://www.theregister.co.uk/2015/12/29/kiev\\_power\\_outages\\_blamed\\_on\\_russian\\_hackers/](https://www.theregister.co.uk/2015/12/29/kiev_power_outages_blamed_on_russian_hackers/).

Leyden, J. (2016) 'Energy firm points to hackers after Kiev power outage', *The Register*. Available at: [https://www.theregister.co.uk/2016/12/21/ukraine\\_electricity\\_outage/](https://www.theregister.co.uk/2016/12/21/ukraine_electricity_outage/).

Leyden, J. (2017) 'North Korean hackers allegedly probing US utilities for weaknesses', *The Register*. Available at: [https://www.theregister.co.uk/2017/10/11/dprk\\_hackers\\_probe\\_us\\_utilities/](https://www.theregister.co.uk/2017/10/11/dprk_hackers_probe_us_utilities/).

Liamputtong, P. (2011) *Focus group methodology: principle and practice*. SAGE Publications. Available at: <http://uk.sagepub.com/en-gb/eur/focus-group-methodology/book233200>.

Lincoln, Y. S., Lynham, S. A. and Guba, E. G. (2011) 'Paradigmatic controversies, contradictions, and emerging confluences, revisited', *The Sage handbook of qualitative research*, 4, pp. 97–128.

Loehlin, J. C. (1998) *Latent variable models: an introduction to factor, path, and*

*structural analysis*. Lawrence Erlbaum Associates Publishers.

Lombardo, T. J. (2017) *The reciprocity of perceiver and environment: the evolution of James J. Gibson's ecological psychology*. Routledge.

Lovejoy, B. (2015) *Security firm publishes list of some of the iOS apps infected by XcodeGhost – including Angry Birds 2, electrek.co*. Available at: <https://9to5mac.com/2015/09/21/xcodeghost-infected-apps/> (Accessed: 14 August 2018).

Lyden, J. (2017) 'World+dog had 1.4 BEEEEELLION of its data records exposed last year', *The Register*. Available at: [https://www.theregister.co.uk/2017/03/28/breach\\_bonanza/](https://www.theregister.co.uk/2017/03/28/breach_bonanza/).

Mader, A. and Srinivasan, S. (2005) 'Curriculum development related to information security policies and procedures', in *Proceedings of the 2nd annual conference on Information security curriculum development*. ACM, pp. 49–53.

Mano Ten Napel and Novealthy (2015) *Wearables and quantified self demand security-first design, Conde Nast Digital*. Available at: <https://www.wired.com/insights/2014/10/wearables-security-first-design/> (Accessed: 13 March 2018).

Marcotte, R. (2013) *Cybersecurity lessons from the New York times security breach, DLT Solutions*. Available at: <http://www.dlt.com/blog/2013/02/05/cybersecurity-lessons-ny-times-security-breach/> (Accessed: 13 August 2018).

Markotten, D. G. tom (2002) *User-centered security engineering*. Available at: <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=9C6208EFF3252E8D51592E83C7BDA1F6?doi=10.1.1.5.3682&rep=rep1&type=pdf>.

Marks, A. and Rezgui, Y. (2009) 'A comparative study of information security awareness in higher education based on the concept of design theorizing', in: Wuhan: IEEE, pp. 1–7. Available at: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5302667>.

Maxion, R. A. and Reeder, R. W. (2005) 'Improving user-interface dependability through mitigation of human error', *International Journal of Human-Computer Studies*, 63(1–2), pp. 25–50. doi: 10.1016/j.ijhcs.2005.04.009.

Mazza, R. and Berre, A. (2007) 'Focus group methodology for evaluating information visualization techniques and tools', in *2007 11th International Conference Information Visualization (IV '07)*. Zurich: IEEE, pp. 74–80. doi: 10.1109/IV.2007.51.

Mc Laughlin, G. H. (1969) 'SMOG grading – a new readability formula', *Journal of reading*, 12(8), pp. 639–646. Available at: <http://www.jstor.org/stable/40011226?seq=1#>.

McNamara, D. S., Graesser, A. C., McCarthy, P. M. and Cai, Z. (2014) *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press. Available at: <http://www.cambridge.org/gb/academic/subjects/psychology/educational-psychology/automated-evaluation-text-and-discourse-coh-metrix?format=PB#Y3b3piPCv3ZEIHVj.97>.

Mendes, E., Mosley, N. and Counsell, S. (2003) 'Early web size measures and

effort prediction for web costimation', *Proceedings - International Software Metrics Symposium*, 2003–Janua, pp. 18–29. doi: 10.1109/METRIC.2003.1232452.

Mertens, D. M. (2014) *Research and evaluation in education and psychology: integrating diversity with quantitative, qualitative, and mixed methods*. Sage publications.

Mesmer, H. A. E. (2008) *Tools for matching readers to texts: research-based practices.*, *Tools for matching readers to texts: Research-based practices*. Guilford Press. Available at: <https://www.guilford.com/books/Tools-for-Matching-Readers-to-Texts/Heidi-Anne-Mesmer/9781593855970>.

Milne, S., Orbell, S. and Sheeran, P. (2002) 'Combining motivational and volitional interventions to promote exercise participation: protection motivation theory and implementation intentions', *British Journal of Health Psychology*, 7, pp. 163–184. doi: 10.1348/135910702169420.

Morgan, D. L. (2007) 'Paradigms lost and pragmatism regained: methodological implications of combining qualitative and quantitative methods', *Journal of mixed methods research*. Sage Publications, 1(1), pp. 48–76.

Moustakas, C. (1994) *Phenomenological research methods*. Sage.

Myry, L., Siponen, M., Pahlila, S., Vartiainen, T. and Vance, A. (2009) 'What levels of moral reasoning and values explain adherence to information security rules? an empirical study', *European Journal of Information Systems*, 18(2), pp. 126–139. doi: 10.1057/ejis.2009.10.

Nardi, P. (2015) *Doing survey research: a guide to quantitative methods*. 3rd Editio. New York, USA: Routledge.

Neal, A. and Griffin, M. A. (2002) 'Safety climate and safety behaviour', *Australian journal of management*. SAGE Publications Sage UK: London, England, 27(1\_suppl), pp. 67–75. Available at: <http://journals.sagepub.com/doi/pdf/10.1177/031289620202701S08>.

Neely, P. (2013) *What's the best time and day to send an email?*, *GetResponse*. Available at: <https://blog.getresponse.com/whats-best-time-day-send-email.html> (Accessed: 22 April 2018).

Neuman, W. L. (2013) *Social research methods: qualitative and quantitative approaches*. Pearson education.

Newman, I. and Benz, C. R. (1998) *Qualitative-quantitative research methodology: exploring the interactive continuum*. SIU Press.

Nicastro, F. M. (2007) 'People, processes, and technology: a winning combination', *Information security management handbook*, pp. 389–399.

Nichols, S. (2018) 'OnePlus minus 40,000 credit cards: smartmobe store hacked to siphon payment info to crooks', *The Register*. Available at: [https://www.theregister.co.uk/2018/01/19/oneplus\\_credit\\_card\\_hack/](https://www.theregister.co.uk/2018/01/19/oneplus_credit_card_hack/).

North, M. M., George, R. and North, S. M. (2006) 'Computer security and ethics awareness in university environments', in *Proceedings of the 44th annual southeast regional conference on - ACM-SE 44*. New York, New York, USA: ACM Press, p. 434. doi: 10.1145/1185448.1185544.

O'Bryan, S., Caraway, R. and CISA, C. (2006) 'Critical elements of information



security program success’, 4(26), p. 22. Available at: <http://www.isaca.org/Journal/Past-Issues/2006/Volume-3/Documents/jpdf0603-Critical-Elements-of-Info.pdf>.

O’Connor, A. D. (1993) ‘Successful strategic information systems planning’, *Information Systems Journal*. Wiley Online Library, 3(2), pp. 71–83.

O’Gorman, L. (2003) ‘Comparing passwords, tokens, and biometrics for user authentication’, *Proceedings of the IEEE*, 91(12), pp. 2019–2020. doi: 10.1109/JPROC.2003.819605.

OCR (2014) *Stolen laptops lead to important HIPAA settlements*, [hhs.gov](https://www.hhs.gov/hipaa/for-professionals/compliance-enforcement/examples/concentra-health-services/index.html). Available at: <https://www.hhs.gov/hipaa/for-professionals/compliance-enforcement/examples/concentra-health-services/index.html> (Accessed: 14 August 2018).

Olusegun, O. J. and Ithnin, N. B. (2013) ‘“ People are the answer to security ”:’, *IJCSIS*, 11(8). Available at: <http://arxiv.org/pdf/1309.0188v1>.

Orlowski, A. (2018) ‘Customers reporting credit card fraud after using OnePlus webstore’, *The Register*. Available at: [https://www.theregister.co.uk/2018/01/15/oneplus\\_users\\_report\\_credit\\_card\\_fraud/](https://www.theregister.co.uk/2018/01/15/oneplus_users_report_credit_card_fraud/).

Palmer, M. E., Robinson, C., Patilla, J. C. and Moser, E. P. (2001) ‘Information security policy framework: best practices for security policy in the E-commerce age’, *Information Systems Security*. Taylor & Francis, 10(2), pp. 1–15. doi: 10.1201/1086/43314.10.2.20010506/31399.4.

Paolacci, G. and Chandler, J. (2014) ‘Inside the Turk’, *Current Directions in Psychological Science*, 23(3), pp. 184–188. doi: 10.1177/0963721414531598.

Parker, D. B. (1997) ‘Information security in a Nutshell’, *Information Systems Security*, 6(1), pp. 14–19. doi: 10.1080/10658989709342524.

Patton, M. Q. (1990) *Qualitative evaluation and research methods*. SAGE Publications, inc.

Pauli, D. (2016a) *Fatal flaws in ten pacemakers make for Denial of Life attacks*, *The Register*. Available at: [https://www.theregister.co.uk/2016/12/01/denial\\_of\\_life\\_attacks\\_on\\_pacemakers/](https://www.theregister.co.uk/2016/12/01/denial_of_life_attacks_on_pacemakers/) (Accessed: 21 May 2018).

Pauli, D. (2016b) ‘Ukraine energy utilities attacked again with open source Trojan backdoor’, *The Register*. Available at: [https://www.theregister.co.uk/2016/01/21/ukraine\\_energy\\_utilities\\_attacked\\_again\\_with\\_open\\_source\\_trojan\\_backdoor/](https://www.theregister.co.uk/2016/01/21/ukraine_energy_utilities_attacked_again_with_open_source_trojan_backdoor/).

Pfleeger, C. P. and Pfleeger, S. L. (2006) ‘Security in computing’. Prentice Hall PTR. Available at: <https://dl.acm.org/citation.cfm?id=1177321>.

Phillips, D. C. and Burbules, N. C. (2000) *Postpositivism and educational research*. Rowman & Littlefield.

Pickard, A. J. (2013) *Research methods in information*. Facet publishing.

Ponemon Institute (2017) *Data risk in the third-party ecosystem*. Available at: [https://cdn2.hubspot.net/hubfs/2575983/Ponemon\\_report\\_Final \(1\).pdf](https://cdn2.hubspot.net/hubfs/2575983/Ponemon_report_Final%20(1).pdf).

Posthumus, S. and Von Solms, R. (2004) ‘A framework for the governance of information security’, *Computers & Security*. Elsevier, 23(8), pp. 638–646.

Powell, R. A., Single, H. M. and Lloyd, K. R. (1996) 'Focus groups in mental health research: enhancing the validity of user and provider questionnaires', *International Journal of Social Psychiatry*. Sage Publications Sage CA: Thousand Oaks, CA, 42(3), pp. 193–206.

Puhakainen, P. (2006) *A design theory for information security awareness, Processing*. University of Oulu. Available at: [https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0CC0QFjAAahUKEwif-7\\_Thu3HAhWKwBQKHVuBAYU&url=http://herkules oulu.fi/isbn9514281144/isbn9514281144.pdf&usg=AFQjCNFpnV9xXbniKUe4wEe2hB4GiO3n4g&sig2=SP0GCdb0AwR83I\\_twi5YJQ](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0CC0QFjAAahUKEwif-7_Thu3HAhWKwBQKHVuBAYU&url=http://herkules oulu.fi/isbn9514281144/isbn9514281144.pdf&usg=AFQjCNFpnV9xXbniKUe4wEe2hB4GiO3n4g&sig2=SP0GCdb0AwR83I_twi5YJQ).

PwC (2013) *Annual report 2013*. Available at: <https://www.pwc.co.uk/assets/pdf/annual-report-2013.pdf>.

PwC (2015) *Information security breaches survey*. Available at: <https://www.pwc.co.uk/assets/pdf/2015-isbs-technical-report-blue-digital.pdf>.

QA Ltd (2014) *Information security awareness course*. Available at: [http://www.qafileshare.com/userUploadAreas/MikeDitchburn/QA Information Security \(June 14\)/course/course.htm?&debug=false](http://www.qafileshare.com/userUploadAreas/MikeDitchburn/QA%20Information%20Security%20(June%2014)/course/course.htm?&debug=false) (Accessed: 29 August 2018).

Qualtrics (2017) *Qualtrics*. Available at: <https://www.qualtrics.com/> (Accessed: 1 January 2016).

RAEng (2018) *Improving cybersecurity requires major coordinated effort, say top engineers, Royal Academy of Engineering*. Available at: <https://www.raeng.org.uk/news/news-releases/2018/march/improving-cybersecurity-requires-major-coordinated> (Accessed: 21 May 2018).

Rainer Jr, R. K., Marshall, T. E., Knapp, K. J. and Montgomery, G. H. (2007) 'Do information security professionals and business managers view information security issues differently?', *Information Systems Security*. Taylor & Francis, 16(2), pp. 100–108. Available at: <https://www.tandfonline.com/doi/full/10.1080/10658980701260579>.

Rankin, E. F. and Culhane, J. W. (1969) 'Comparable cloze and multiple-choice comprehension test scores', *Journal of Reading*, 13(3), pp. 193–198. Available at: <http://www.jstor.org/stable/40017267>.

Rastogi, R. (2012) 'Information security service branding – beyond information security awareness', *Journal of Systemics, Cybernetics & Informatics*, 10(6), pp. 54–59. Available at: [http://www.iiisci.org/Journal/CV\\$/sci/pdfs/HEA522ET.pdf](http://www.iiisci.org/Journal/CV$/sci/pdfs/HEA522ET.pdf).

Readabilityformulas.com (2017) *The SPACHE readability formula for young readers*. Available at: <http://www.readabilityformulas.com/spache-readability-formula.php> (Accessed: 6 November 2017).

ReadabilityFormulas.com (2017) *The automated readability index (ARI), ReadabilityFormulas.com*. Available at: <http://www.readabilityformulas.com/automated-readability-index.php> (Accessed: 5 November 2017).

Reber, A. S. (1995) *The Penguin dictionary of psychology*. Penguin Press.

Redish, J. (2000) 'Readability formulas have even more limitations than Klare

discusses', *ACM Journal of Computer Documentation*, 24(3), pp. 132–137. doi: 10.1145/344599.344637.

Rezgui, Y. and Marks, A. (2008) 'Information security awareness in higher education: an exploratory study', *Computers & Security*. Elsevier Ltd, 27(7–8), pp. 241–253. doi: 10.1016/j.cose.2008.07.008.

Rezmierski, V. E., Seese Jr, M. R. and Clair II, N. S. (2002) 'University systems security logging: who is doing it and how far can they go?', *Computers & Security*. Elsevier, 21(6), pp. 557–564.

Richards, J. C. and Schmidt, R. (2015) *Longman dictionary of language teaching and applied linguistics*. Routledge.

Roper, C., Fischer, L. and Grau, J. (2005) *Security education, awareness and training: seat from theory to practice*. Butterworth Heinemann.

Rossmann, G. B. and Wilson, B. L. (1985) 'Numbers and words: combining quantitative and qualitative methods in a single large-scale evaluation study', *Evaluation review*. Sage Publications Sage CA: Thousand Oaks, CA, 9(5), pp. 627–643.

Rouse, M. (2005) *Dictionary attack*, *WhatIs.com*. Available at: <http://searchsecurity.techtarget.com/definition/dictionary-attack> (Accessed: 9 October 2014).

SANS (2010) *Securing the human*. Available at: <https://www.sans.org/security-awareness-training/blog/securing-human> (Accessed: 29 August 2018).

Sapsford, R. and Jupp, V. (2006) *Data collection and analysis*. Sage.

Sarantakos, S. (2012) *Social research*. Palgrave Macmillan.

Sasse, M. A., Brostoff, S. and Weirich, D. (2001) 'Transforming the "Weakest Link" — a human/computer interaction approach to usable and effective security', *BT Technology Journal*. Kluwer Academic Publishers, 19(3), pp. 122–131. doi: 10.1023/A:1011902718709.

Sasse, M. A. and Flechais, I. (2005) 'Usable security: why do we need it? how do we get it?', in O'Reilly. Available at: <http://discovery.ucl.ac.uk/20345/2/cransimpsonbook.pdf>.

Saunders, M., Lewis, P. and Thornhill, A. (2009) *Research methods for business students*, TBS. doi: 10.1017/CBO9781107415324.004.

Savin-Baden, M. and Major, C. H. (2013) *Qualitative research: the essential guide to theory and practice*. Routledge.

Scandura, T. A. and Williams, E. A. (2000) 'Research methodology in management: current practices, trends, and implications for future research', *Academy of Management journal*. Academy of Management, 43(6), pp. 1248–1264.

Scheaffer, R. L., Mendenhall III, W., Ott, R. L. and Gerow, K. G. (2011) *Elementary survey sampling*. 7th edn. Duxbury Press. Available at: <https://www.amazon.com/Elementary-Survey-Sampling-Richard-Scheaffer/dp/0840053614>.

Schmidt, G. B. (2015) 'Fifty days an MTurk worker: the social and motivational context for Amazon mechanical turk workers', *Industrial and Organizational Psychology*, 8(02), pp. 165–171. doi: 10.1017/iop.2015.20.

Schneier, B. (2001) 'Managed security monitoring: network security for the 21st century', *Computers & Security*. Elsevier Advanced Technology, 20(6), pp. 491–503. Available at: <https://www.schneier.com/academic/paperfiles/paper-msm.pdf>.

Schneier, B. (2004) *Secrets and lies: digital security in a networked world*. Wiley computer Publishing.

Schneier, B. (2006) *Beyond fear: thinking sensibly about security in an uncertain world*. Springer Science & Business Media.

Schou, C. D. and Trimmer, K. J. (2004) 'Information assurance and security', *Journal of Organizational and End User Computing*. IGI Global, 16(3), pp. 123–145. Available at: [https://www.researchgate.net/profile/Sugata\\_Sanyal/publication/220349069\\_Information\\_Assurance\\_and\\_Security/links/0deec52a0a73ee8159000000.pdf](https://www.researchgate.net/profile/Sugata_Sanyal/publication/220349069_Information_Assurance_and_Security/links/0deec52a0a73ee8159000000.pdf).

Schumacker, R. E. and Lomax, R. G. (2012) *A beginner's guide to structural equation modeling*. Routledge.

Schutt, R. K. (2011) *Investigating the social world: the process and practice of research*. Pine Forge Press.

Schwandt, T. A. (2000) 'Three epistemological stances for qualitative inquiry: interpretivism, hermeneutics, and social constructionism', *Handbook of qualitative research*, 2, pp. 189–213.

Schwarm, S. E. and Ostendorf, M. (2005) 'Reading level assessment using support vector machines and statistical language models', in *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*. Association for Computational Linguistics, pp. 523–530.

Senter, R. J. and Smith, E. A. (1967) *Automated readability index*. CINCINNATI UNIV OH. Available at: <http://www.dtic.mil/dtic/tr/fulltext/u2/667273.pdf>.

Shaw, R. S., Chen, C. C., Harris, A. L. and Huang, H.-J. (2009) 'The impact of information richness on information security awareness training effectiveness', *Computers & Education*, 52(1), pp. 92–100. doi: 10.1016/j.compedu.2008.06.011.

Si, L. and Callan, J. (2001) 'A statistical model for scientific readability', in *Proceedings of the tenth international conference on Information and knowledge management*. ACM, pp. 574–576.

Siegel, C. A., Sagalow, T. R. and Serritella, P. (2002) 'Cyber-risk management: technical and insurance controls for enterprise-level security', *Information Systems Security*. Taylor & Francis, 11(4), pp. 33–49. Available at: <http://www.eprivacy.com/lectures/cyber-risk.pdf>.

Singer, H. and Donlan, D. (1982) 'Active comprehension: problem-solving schema with question generation for comprehension of complex short stories', *Reading Research Quarterly*. JSTOR, pp. 166–186.

Siponen, M., Adam Mahmood, M. and Pahlila, S. (2014) 'Employees' adherence to information security policies: an exploratory field study', *Information and Management*. Elsevier B.V., 51(2), pp. 217–224. doi: 10.1016/j.im.2013.08.006.

Siponen, M., Pahlila, S. and Mahmood, M. A. (2010) 'Compliance with information security policies: an empirical investigation', *Computer*. IEEE, 43(2).

Siponen, M. T. and Oinas-kukkonen, H. (2007) *A review of information security*

issues and respective contributions. Available at: [http://dl.acm.org/ft\\_gateway.cfm?id=1216224&type=pdf&CFID=581787190&CFTOKEN=90143189](http://dl.acm.org/ft_gateway.cfm?id=1216224&type=pdf&CFID=581787190&CFTOKEN=90143189).

Siponen, M. and Willison, R. (2009) 'Information security management standards: problems and solutions', *Information & Management*. Elsevier, 46(5), pp. 267–270.

Slattery, T. J. and Rayner, K. (2009) 'The influence of text legibility on eye movements during reading', *Applied Cognitive Psychology*, 24(8), pp. 1129–1148. doi: 10.1002/acp.1623.

Sohrabi Safa, N., Von Solms, R. and Furnell, S. (2016) 'Information security policy compliance model in organizations', *Computers & Security*. Elsevier Advanced Technology, 56, pp. 70–82. doi: 10.1016/J.COSE.2015.10.006.

Von Solms, R. (1999) 'Information security management: why standards are important', *Information Management & Computer Security*. MCB UP Ltd, 7(1), pp. 50–58. Available at: [http://130.18.86.27/faculty/warkentin/SecurityPapers/Robert/Others/VonSolms1999\\_IMCS7\\_1\\_InfoSecStandards.pdf](http://130.18.86.27/faculty/warkentin/SecurityPapers/Robert/Others/VonSolms1999_IMCS7_1_InfoSecStandards.pdf).

Von Solms, R. and Von Solms, B. (2004) 'From policies to culture', *Computers & Security*, 23(4), pp. 275–279. doi: 10.1016/j.cose.2004.01.013.

Sommestad, T., Hallberg, J., Lundholm, K. and Bengtsson, J. (2014) 'Variables influencing information security policy compliance', *Information Management & Computer Security*, 22(1), pp. 42–75. doi: 10.1108/IMCS-08-2012-0045.

Sonderegger, A. and Sauer, J. (2010) 'The influence of design aesthetics in usability testing: effects on user performance and perceived usability', *Applied Ergonomics*, 41(3), pp. 403–410. doi: 10.1016/j.apergo.2009.09.002.

Spero, J. (2018) *Hackers steal financial data in BA website attack*, *Financial Times*. Available at: <https://www.ft.com/content/3d569654-b1fe-11e8-8d14-6f049d06439c> (Accessed: 9 September 2018).

Stake, R. E. (1995) *The art of case study research*. Sage.

Sterling, B. (2012) *Spear-phishing and water-holing*, *Conde Nast Digital*. Available at: <https://www.wired.com/2012/10/spear-phishing-and-water-holing/> (Accessed: 13 August 2018).

Stoneburner, G., Goguen, A. Y. and Feringa, A. (2002) 'Risk management guide for information technology systems'. National Institute of Standards & Technology. Available at: Risk management guide for information technology systems.

Straub, D. W. (1990) 'Effective IS security: an empirical study', *Information System Research*. INFORMS, 1(3), pp. 255–276. doi: 10.1287/isre.1.3.255.

Strauss, A. L. and Corbin, J. (1998) 'Basics of qualitative research: techniques and procedures for developing grounded theory', *Newbury Park CA*.

Sumeeth, M., Singh, R. I. and Miller, J. (2010) 'Are online privacy policies readable?', *International Journal of Information Security and Privacy*. Hershey, PA, USA: IGI Global, 4(1), pp. 93–116. doi: 10.4018/jisp.2010010105.

Suo, X., Zhu, Y. and Owen, G. S. (2005) 'Graphical passwords: a survey', in *Proceedings - Annual Computer Security Applications Conference, ACSAC*. Tucson: IEEE, pp. 463–472. doi: 10.1109/CSAC.2005.27.

Taipale, K. A. (2004) 'Technology, security and privacy: the fear of frankenstein, the mythology of privacy and the lessons of King Ludd', *Yale JL & Tech*. HeinOnline, 7, p. 123. Available at: <http://information-retrieval.info/papers/TSP-YLS.pdf>.

Taylor, W. (1953) "'Cloze procedure": a new tool for measuring readability', *Journalism Quarterly*, 30, pp. 415–433. Available at: <http://psycnet.apa.org/psycinfo/1955-00850-001>.

Teddlie, C. and Tashakkori, A. (2010) *Handbook of mixed methods in social & behavioral research*. Sage.

Terrell, S. R. (2016) 'Writing a proposal for your dissertation: guidelines and examples', in. Guilford Press, p. 282. Available at: <https://www.amazon.co.uk/Writing-Proposal-Your-Dissertation-Guidelines/dp/1462523021>.

Thomas, B. (2005) *Simple formula for strong passwords (SFSP) tutorial*. Available at: <https://www.sans.org/reading-room/whitepapers/authentication/simpleformula-strong-passwords-sfsp-tutorial-1636>.

Thomas, D. R., Beresford, A. R. and Rice, A. (2015) 'Security metrics for the android ecosystem', in *Proceedings of the 5th Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices*. ACM, pp. 87–98.

Thomas, M. (2004) *Network security first-step*. Cisco Press.

Thomson, I. (2018) 'Google's socially awkward geeks craft socially awkward AI bot that calls people for you', *The Register*.

Thomson, K. and von Solms, R. (2007) 'Cultivating corporate information security obedience'. Retrieved March. Available at: <https://pdfs.semanticscholar.org/8759/0ed6f4605050b863f4ea7c2d77be4711dfe8.pdf>.

Tong, F. (2018) 'Online retail sales in China soar past \$1 trillion in 2017', *DigitalCommerce360*. Available at: <https://www.digitalcommerce360.com/2018/02/08/online-retail-sales-china-soar-past-1-trillion-2017/>.

Trevino, L. K. (1992) 'The social effects of punishment in organizations: a justice perspective', *Academy of Management Review*, 17(4), pp. 647–676. doi: 10.5465/AMR.1992.4279054.

Tryfonas, T., Kiountouzis, E. and Poulymenakou, A. (2001) 'Embedding security practices in contemporary information systems development approaches', *Information Management & Computer Security*. MCB UP Ltd, 9(4), pp. 183–197.

Tsohou, A., Kokolakis, S., Karyda, M. and Kiountouzis, E. (2008) 'Investigating information security awareness: research and practice gaps', *Information Security Journal: A Global Perspective*, 17(5–6), pp. 207–227. doi: 10.1080/19393550802492487.

Tudor, J. (2000) *Information security architecture: an integrated approach to security in the organization*. Boca Raton FL: CRC Press.

Vance, A. (2010) *Why do employees violate is security policies? insights from multiple theoretical perspectives*, October. University of Oulu. Available at:

[https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=0C CQQFjABahUKEwiRkLe\\_9ezHAAhXBQBQKHcj6AWw&url=http://herkules.oulu.fi/isbn9789514262876/isbn9789514262876.pdf&usg=AFQjCNGVAvSPS2JLLy5Hy2ZzfOrhfY3XSQ&sig2=wMVw23RoAq\\_OPambty79fg&cad=rja](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=0C CQQFjABahUKEwiRkLe_9ezHAAhXBQBQKHcj6AWw&url=http://herkules.oulu.fi/isbn9789514262876/isbn9789514262876.pdf&usg=AFQjCNGVAvSPS2JLLy5Hy2ZzfOrhfY3XSQ&sig2=wMVw23RoAq_OPambty79fg&cad=rja).

Vanderkam, L. (2012) *The best time to send email so it will get read*, CBS News. Available at: <http://www.cbsnews.com/news/the-best-time-to-send-email-so-it-will-get-read/> (Accessed: 12 February 2016).

De Vaus, D. (2013) *Surveys in social research (social research today)*. Routledge.

VerticalResponse (2017) *The surprisingly best times to send your email marketing campaigns*. Available at: <https://www.verticalresponse.com/blog/best-time-send-email-marketing-campaign/> (Accessed: 22 April 2018).

Vroom, C. and Von Solms, R. (2004) 'Towards information security behavioural compliance', *Computers & Security*, 23(3), pp. 191–198. doi: 10.1016/j.cose.2004.01.012.

Wagner, A. E. and Brooke, C. (2007) 'Wasting time: the mission impossible with respect to technology-oriented security approaches', *The Electronic Journal of Business Research Methods*, 5(2), pp. 117–124. Available at: <http://www.ejbrm.com/issue/download.html?idArticle=177>.

Waly, N., Tassabehji, R. and Kamala, M. A. (2012) 'Measures for improving information security management in organisations: the impact of training and awareness programmes', in *UKAIS*, p. 8. Available at: <https://pdfs.semanticscholar.org/207e/648e5c42721817b0d10b4dc5e600acc2c1ad.pdf>.

Wearden, G. (2007) *U.K. company fined over laptop theft*, cent.com. Available at: <https://www.cnet.com/news/u-k-company-fined-over-laptop-theft/> (Accessed: 6 August 2018).

Weir, G. and Anagnostou, N. (2008) 'Collocation frequency as a readability factor', *Proceedings of the 13th Conference of the Pan Pacific Association of Applied Linguistics*. Available at: [http://www.cis.strath.ac.uk/cis/research/publications/papers/strath\\_cis\\_publication\\_2261.pdf](http://www.cis.strath.ac.uk/cis/research/publications/papers/strath_cis_publication_2261.pdf).

Weir, G. and Ritchie, C. (2006) 'Estimating readability with the Strathclyde readability measure', *ICT in the Analysis, Teaching and Learning of Languages, Preprints of the ICTATLL Workshop 2006*, pp. 25–32. Available at: <http://strathprints.strath.ac.uk/2380/1/strathprints002380.pdf> (Accessed: 5 March 2015).

Whitman, M. E. and Mattord, H. J. (2011) *Principles of information security*. Cengage Learning. Available at: [https://college.cengage.com/information\\_security/course360/principles\\_of\\_information\\_security\\_1133135005/ebook/whitman\\_1111138214\\_ch10.pdf](https://college.cengage.com/information_security/course360/principles_of_information_security_1133135005/ebook/whitman_1111138214_ch10.pdf).

Wiant, T. L. (2005) 'Information security policy's impact on reporting security incidents', *Computers & Security*, 24(6), pp. 448–459. doi: 10.1016/j.cose.2005.03.008.

Wilson, K. and Wauson, J. (2010) *The AMA handbook of business writing: the ultimate guide to style, grammar, punctuation, usage, construction, and formatting*,

AMACOM. AMACOM. doi: 808/.06665.

Wilson, M. and Hash, J. (2003) *Building an information technology security awareness and training Program. NIST special publication 800-50*. Available at: <http://csrc.nist.gov/publications/nistpubs/800-50/NIST-SP800-50.pdf>.

Wilson, W. M., Rosenberg, L. H. and Hyatt, L. E. (1997) 'Automated analysis of requirement specifications', in *Proceedings of the 19th International Conference on Software Engineering*. New York, NY, USA: ACM (ICSE '97), pp. 161–171. doi: 10.1145/253228.253258.

Wolcott, H. F. (2008) *Ethnography: a way of seeing*. AltaMira Press.

Wombat security (2017) *Try our interactive training modules*. Available at: [https://www.wombatsecurity.com/try-our-interactive-training-modules?utm\\_term=wombat\\_security&utm\\_campaign=Branded++UK&utm\\_source=adwords&utm\\_medium=ppc&hsa\\_acc=8253056476&hsa\\_net=adwords&hsa\\_cam=741275167&hsa\\_ad=243213335940&hsa\\_kw=wombat\\_security&hsa\\_gr](https://www.wombatsecurity.com/try-our-interactive-training-modules?utm_term=wombat_security&utm_campaign=Branded++UK&utm_source=adwords&utm_medium=ppc&hsa_acc=8253056476&hsa_net=adwords&hsa_cam=741275167&hsa_ad=243213335940&hsa_kw=wombat_security&hsa_gr) (Accessed: 29 August 2018).

Wood, C. C. (1982) 'Policies for deterring computer abuse', *Computers & Security*, 1(2), pp. 139–145. doi: 10.1016/0167-4048(82)90006-2.

Wood, M. B. (1984) *Introducing computer security*. Bookmen Associates.

Workman, M., Bommer, W. H. and Straub, D. (2008) 'Security lapses and the omission of information security measures: a threat control model and empirical test', *Computers in human behavior*. Elsevier, 24(6), pp. 2799–2816. Available at: <https://www.sciencedirect.com/science/article/pii/S0747563208000824/pdf?md5=bea5b53823ed1458866964df3d736cfd&pid=1-s2.0-S0747563208000824-main.pdf>.

Yan, J., Blackwell, A., Anderson, R. and Grant, A. (2004) 'Password memorability and security: empirical results', *IEEE Security & Privacy Magazine*, 2(5), pp. 25–31. doi: 10.1109/MSP.2004.81.

Yin, R. K. (2017) *Case study research and applications: design and methods*. Sage publications.

Yunos, Z., Hamid, R. S. A. and Ahmad, M. (2016) 'Development of a cyber security awareness strategy using focus group discussion', in *2016 SAI Computing Conference (SAI)*. IEEE, pp. 1063–1067. doi: 10.1109/SAI.2016.7556109.

Zhou, S., Jeong, H. and Green, P. A. (2017) 'How consistent are the best-known readability equations in estimating the readability of design standards?', *IEEE Transactions on Professional Communication*, 60(1), pp. 97–111. doi: 10.1109/TPC.2016.2635720.

Zikmund, W. G., Babin, B. J., Carr, J. C. and Griffin, M. (2013) *Business research methods, Business Research Methods*. doi: 10.1.1.131.2694.

Zurko, M. E. (2005) 'User-centered security: stepping up to the grand challenge', *21st Annual Computer Security Applications Conference (ACSAC'05)*. Ieee, (Acsac), pp. 187–202. doi: 10.1109/CSAC.2005.60.

Zurko, M. E., Kaufman, C., Spanbauer, K. and Bassett, C. (2002) 'Did you ever have to make up your mind? what notes users do when faced with a security decision', in *Proceedings of 18th Annual Computer Security Applications Conference*. IEEE Comput. Soc, pp. 371–381. doi: 10.1109/CSAC.2002.1176309.



## APPENDIX (A)

<sup>2</sup> FRE Scale		<sup>3</sup> SRM Scale	
Mark	Readability Category	Mark	Readability Category
0-29	Very Confusing	< 30	Very Confusing
30-49	Difficult	30-40	Difficult
50-59	Fairly difficult	40-50	Fairly difficult
60-69	Standard	50-65	Standard
70-79	Fairly easy	65-80	Easy
80-89	Easy	> 80	Very easy
90-100	Very easy		

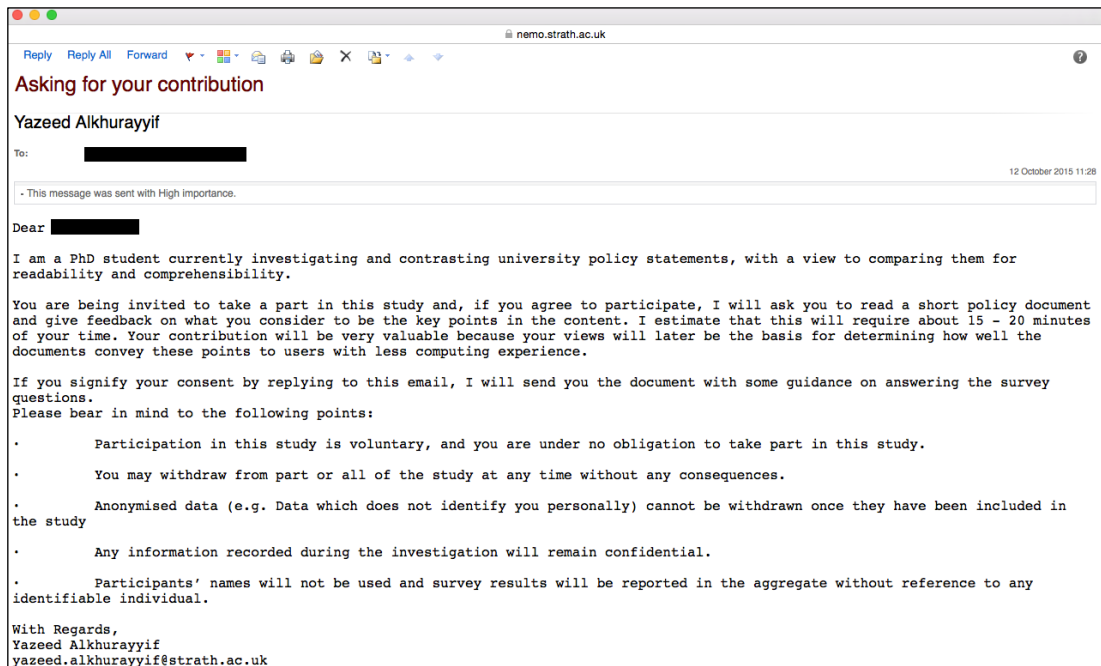
Table A3. 1: Estimate of readability on the FRE and SRM scale

---

<sup>2</sup> Source: Adopted from Flesch, R., 1974. *The Art of Readable Writing*, New York, USA: Harper and Row cited by Ammann, F. & Sowa, A., 2013. *Readability as Lever for Employees' Compliance With Information Security Policies*. ISACA, 4, pp.1-4.

<sup>3</sup> Source: Adopted from Weir, G. R., & Ritchie, C. 2006. *Estimating readability with the Strathclyde readability measure*. In *ICT in the Analysis, Teaching and Learning of Languages*, Preprints of the ICTATLL Workshop 2006 (pp. 25-32).

## APPENDIX (B)



Screenshot 5.1: The first introductory email





Screenshot 5.2: The second introductory email

## Ethical Approvals:

### Ethics application has been approved

---

www-data [www-data@cis.strath.ac.uk]    Actions

To: Yazeed Alkhurayyif

*Inbox* 14 April 2015 13:32

Hello,




Your ethics application "Investigating the Influence of Readability on Information Security Policies." (ID: 264) has been approved.

URL: <https://local.cis.strath.ac.uk/local/ethics/index.php?view=264>

Ethics Approval System.

### Ethics application has been approved

---

www-data [www-data@cis.strath.ac.uk]    Actions

To: Yazeed Alkhurayyif

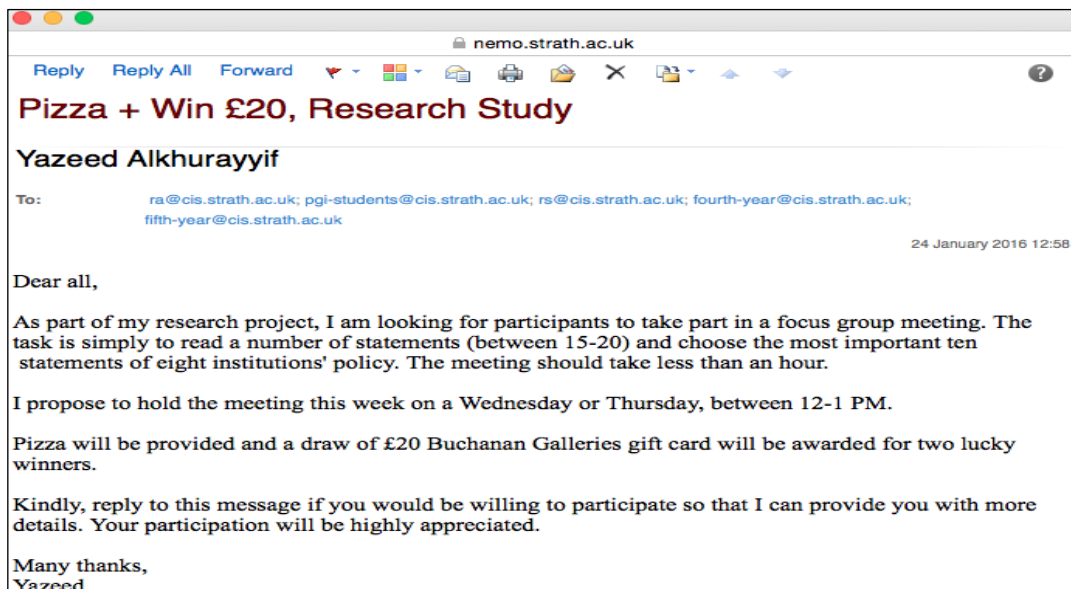
*Inbox* 28 April 2016 12:32

Hello,

Your ethics application "Investigating the Influence of Readability on Information Security Policies 2." (ID: 418) has been approved.

URL: <https://local.cis.strath.ac.uk/wp/extras/ethics/index.php?view=418>

Ethics Approval System.



Screenshot 5.3: The email to recruit focus group participants

	<b>Count</b>	<b>%</b>
<b>Gender</b>		
Male	15	88.2
Female	2	11.8
<b>Age</b>		
26-34	13	76.5
35-54	3	17.6
55 and above	1	5.9
<b>Qualification</b>		
PhD	2	11.8
MSc	12	70.6
BSc	3	17.6
<b>Computer Experience</b>		
Proficient	7	41.2
Intermediate	8	47.1
Basic	2	11.8
<b>Study Subject</b>		
Art, design, media & performance	1	5.9
Biological, health, chemical and agricultural sciences	2	11.8
Business, law, logistics and management	1	5.9
Computing and mathematical sciences	3	17.6
Education and teaching	1	5.9
Engineering and robotics	3	17.6
Hospitality, tourism, sport and leisure	1	5.9
Humanities and languages	2	11.8
Marine, earth, geography and environment	1	5.9
Others	2	11.8

Table A5. 1: The second pilot study's demographic information

	Participants' count	Mean*
<b>Group</b>		
First group	9	55.5556
Second group	8	67.8125
<b>Qualification</b>		
PhD	2	65.0000
MSc	12	61.8750
BSc	3	30.0000
<b>Computer Experience</b>		
Proficient	7	64.2857
Intermediate	8	66.5625
Basic	2	30.0000
*Converted into percentages		

Table A5. 2: The second pilot study's participants results

	Policy A	Policy B	Policy C	Policy D	Policy E	Policy F	Policy G	Policy H
Participants' count	9	9	9	9	8	8	8	8
Mean	6.111	6.222	5.555	4.333	7.375	7.125	7.000	5.625
Mode	7.00	8.00	8.00	4.00	8.00	5, 9 & 8*	6, 7 & 8*	6.00
Std. Deviation	1.269	1.986	2.697	2.397	2.445	1.642	1.772 8	1.5979
Variance	1.161	3.944	7.278	5.750	5.982	2.696	3.143	2.554
*. Multiple modes								

Table A5. 3: The second pilot study's descriptive statistics

## Exploring how readily statements in security policies can be understood

To participate in this study, your email is required in order to send you a unique username and password. when you submit your email, you will receive an email titled "Your [learnclick.com](https://learnclick.com) account" that has a username, password and the study link. For your comfort advised, it is easier to use a computer rather than a mobile to answer the study Thank you

\* Required

**E-mail \***

Your answer \_\_\_\_\_

**Surname (option)**

Your answer \_\_\_\_\_

**SUBMIT**

Never submit passwords through Google Forms.

Screenshot 5.4: Google form for the second pilot study

Drag the boxes onto the matching gaps.

information	legitimate	eligibility	unsolicited	encouraged	violates	resources
fraudulent	violated	unlawful				

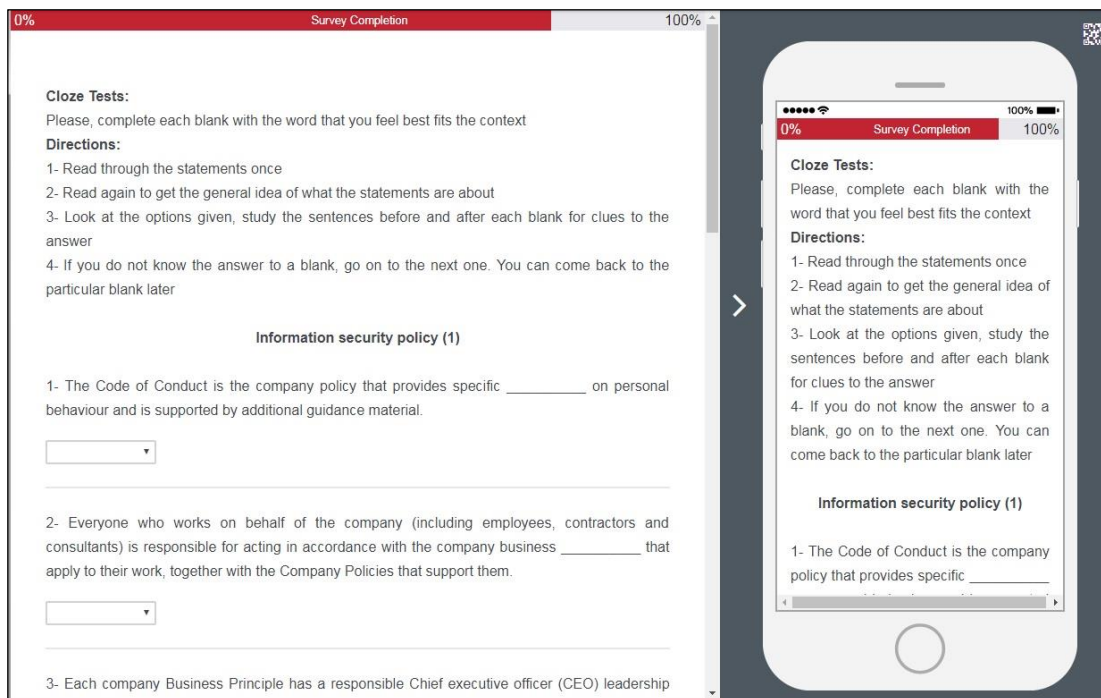
### Information security policy (2)

- Sharing an online identity (user ID and password or other authenticator such as a token or certificate) \_\_\_\_\_ University policy.
- Users must not send, view or download \_\_\_\_\_, harassing, obscene (i.e., pornographic), threatening, or other messages or material that are a violation of applicable law or University policy.
- \_\_\_\_\_ file-sharing using the University's information resources is a violation of copyrights and licenses policy.
- A user found to have \_\_\_\_\_ Information Resources policy will be subject to appropriate disciplinary action up to and including discharge, dismissal, expulsion, and/or legal action.
- Authorized system administrators may access information resources, but only for a \_\_\_\_\_ operational purpose and only the minimum access required to accomplish this legitimate operational purpose.
- Any personally owned \_\_\_\_\_ used for University business are subject to this policy and must comply with all the University requirements pertaining to that type of resource and to the type of data involved.
- Users must not encroach, disrupt or otherwise interfere with access or use of the University's information or information resources. For the avoidance of doubt, without express permission, users must not give away University information or send bulk \_\_\_\_\_ email.
- All University \_\_\_\_\_ is classified into one of 4 levels based on sensitivity and risk.
- \_\_\_\_\_ ends when a person's active association with the University ends.
- Users are strongly \_\_\_\_\_ to change their password regularly (at least once every three months).

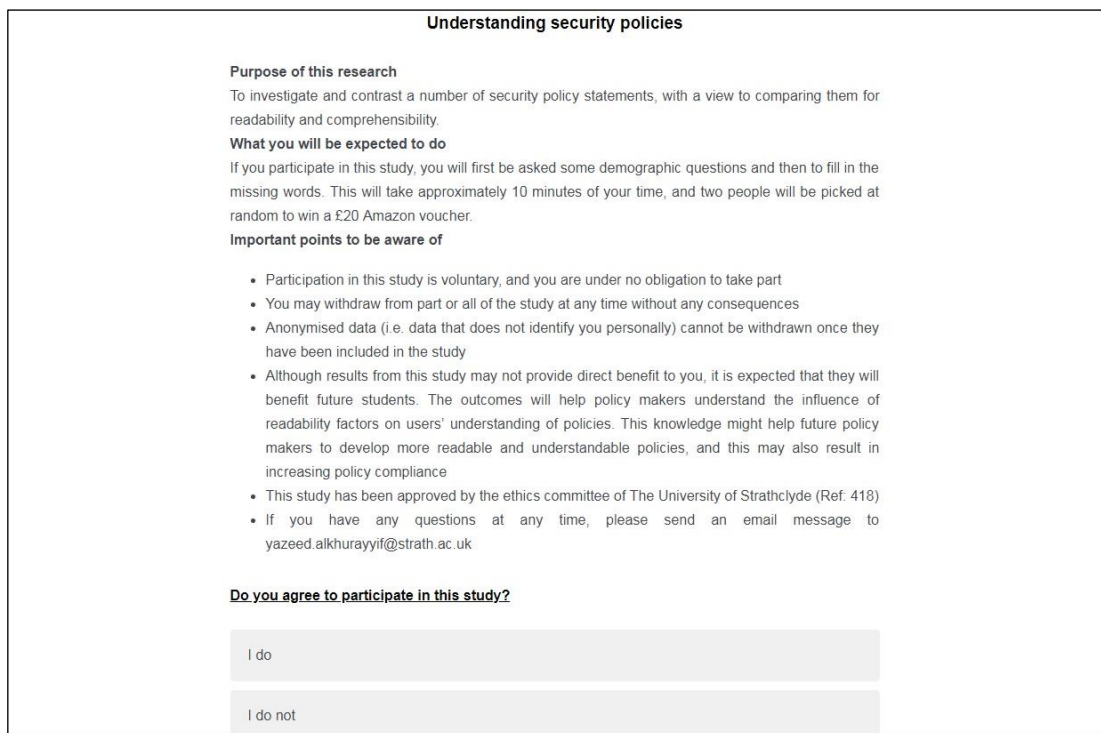
1 attempts remaining

**Submit**

Screenshot 5.5: The second pilot study's Cloze test example



Screenshot 5.6: The main study's Cloze test example



Screenshot 5.7: The main study's cover letter



## APPENDIX (C)

Text	Human Mean	Rank		SRM1 Scale
		Human	SRM1	
Policy D “Imperial University”	5.0851	1	1	40.44
Policy C “Melbourne University”	5.9043	2	2	42.65
Policy F “Stanford University”	6.0577	3	3	43.39
Policy B “Princeton University”	6.2165	4	4	43.77
Policy A “TELUS Company”	6.5052	5	5	46.15
Policy E “Telstra Company”	7.0962	6	7	62.46
Policy H “Cambridge University”	7.1584	7	6	61.63
Policy G “T-Mobile Company”	7.9208	8	8	63.77

Table A6. 1: Comparison of human results VS SRM 1 results