Readability as a Basis for Information Security Policy Assessment

Yazeed Alkhurayyif
Department of Computer and Information Sciences
University of Strathclyde
Glasgow, UK

George R S Weir
Department of Computer and Information Sciences
University of Strathclyde
Glasgow, UK

***Abstract*— Most organisations now impose information security policies (ISPs) or 'conditions of use' agreements upon their employees. The need to ensure that employees are informed and aware of their obligations toward information security is apparent. Less apparent is the correlation between the provision of such policies and their compliance.**

In this paper, we report our research into the factors that determine the efficacy of information security policies (ISPs). Policies should comprise rules or principles that users can easily understand and 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.

The present study investigates the effectiveness of applying readability metrics as an indicator of policy comprehensibility. Results from a preliminary study reveal variations in the comprehension test results attributable to the difficulty of the examined policies. The pilot study shows some correlation between the software readability formula results and human comprehension test results and supports our view that readability has an impact upon understanding ISPs.

These findings have important implications for users' compliance with information security policies and suggest that the application of suitably selected readability metrics may allow policy designers to evaluate their draft policies for ease of comprehension prior to policy release. Indeed, there may be grounds for a readability compliance test that future ISPs must satisfy.

Keywords— Information security policy, ISP, Information security, Readability, ISP compliance.

I. INTRODUCTION

With advances in technology and the increase in its use across organisations, the demand to protect confidential information from prying eyes has become the requirement of the age. This need has given birth to information security awareness (ISA) programmes in corporations worldwide. Many studies indicate that employee attitudes and lack of security awareness are the most notable contributors to security incidents [1]. Institutions should have a security policy that includes pertinent documentation that reflects an organisation's IS philosophy and commitment [2]. As Higgins [3] emphasises "without a policy, security practices will be developed without clear demarcation of objectives and responsibility". The ambition of setting up an effective ISP will not be achieved unless users are familiar with its content and comply with its requirements. Therefore, institutions should strive to achieve information security policy compliance via comprehensibility. In light of changing circumstances and technological progress, ISPs should be regularly enhanced and updated to maintain fit to the institution's vision and mission.

Several factors enable compliance with regulations and rules of security and one of these is the comprehensibility of the ISP itself. In part, this can be estimated by applying a readability formula to the text of information security policies. Readability is a characterisation of how straightforwardly textual material can be read and understood and over the years many different readability formulae have been proposed. For example, McLaughlin [4] developed a readability formula as a mathematical equation influenced by regression analysis, which shows the relationship between two variables, a gauge of the difficulty experienced by individuals reading a given written material, and a measure of the linguistic characteristics of that content. Such a formula can be utilised to estimate reading difficulty from the linguistic characteristics of the texts.

In what follows, we outline pertinent literature before detailing the methodology applied in this research. This includes our selection of eight sample policies, an experts' insight stage, focus group interviews, development of comprehension tests and a pilot study, before making a comparison of comprehension results against a specially selected readability metric. We detail the data collection and analysis and discuss the results of the pilot study using readability metrics. The final section provides a summary and suggestions for future work.

II. BACKGROUND

A. Prior Research Related to the Important of ISA Programmes

Information is one of the most valuable commodities in today's world. 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 gains. 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 into the system [5].

The primary aim of information security awareness 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 legal responsibility to ensure that their confidential information is protected from malicious access [6].

Studies indicate that employees are the weakest links in the information security chain, which is why information security awareness (ISA) is considered important [7]. One way to reduce the incidence and severity of incidents is to raise the level of information security awareness within organisations and the public [5]. Although information security policies 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.

To ensure minimum losses and the safety of online data, institutions are working to make information security awareness a primary concern. Particularly in the financial and banking sector, organisations present employees with guidelines for the protection of corporate data. General security awareness, organisational budgets and level of employee computer skills are noted as major barriers to increasing information security awareness [8].

B. The Role and Importance of Information Security Policy

Information security policies often form part of an organisation's official regulatory framework. The role of the information security policy is to ensure that any decisions and actions are consonant with the objectives of an organisation.

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 a accepted as part of regular tasks [9]. Users are frequently identified as the key vulnerability to an organisation's information security and are often the main causes of security incidents. Höne and Eloff [10] believe that users ignore ISPs because they do not fully understand the policy. Hence, the main cause of security incidents may be the security policy itself, if the users do not fully understand its contents. Accordingly, authors who are responsible for writing an ISP should seek to ensure that the information in the policy reaches its audience easily and effectively.

C. How to Make Successful ISPs

There is no obvious single approach to ISP design and content that is guaranteed to help an organisation accomplish its information security aims, but key among the requirements is successfully clarifying the requirements and concepts of the information security policy to the users (op. cit.). Thereby, evaluating policy comprehensibility can assist in determining whether the ISP is likely to be effective or not. For instance, if the auditors of an ISP certify that controls and security measures are working sufficiently with the policy, this indicates a good fit between the policy and those charged with its application. Of course, in the contrary situation, insufficient controls and security measures may produce an ineffective policy [9].

Authors should consider the writing style and the way in which the ISP is presented to users. Höne and Eloff [10] suggest that the ISP document should be presented in beautiful and attractive style in order to catch the users' attention and ensure the desired objectives are delivered. Notably, organisations should not leave the documenting of ISPs to technical staff in isolation from others. Although they may have experience of information security technologies, this may not be matched in experience of users' understanding and how IS may suit the broader organisational culture (op. cit.).

An ISP cannot be successful unless users are familiar with it. Consequently, institutions should strive to distribute the ISP efficiently and be approachable to address any issues related to the transparency of its content. Additionally, the ISP should be regularly enhanced and updated to ensure continued 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.

D. Gauging the Success or Effectiveness of ISPs

Some factors may minimise the efficacy of ISPs even before the ISP is introduced (proactive/prior factors) while other factors may minimise the efficacy after the ISP is in 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 on information asset protection, achieve ISPs target within available budget, balance of effort between achieving short-term goals with anticipating long-term targets, extent of alignment of ISPs with the organisation's objectives and cost justification for IS [11]–[14].

The present paper focuses on assessing the readability factor in affecting the success or effective operation of ISPs. The ease of reading ISP documents is a "proactive factor". Investigating the readability of information security policies may help in complying with regulations and rules of security, which result in increased effectiveness of ISPs. In principle, this may be achieved by testing documents using software readability metrics and/or testing human readers using comprehension tests. Readability metrics may help in assessing the quality of ISPs and could offer an easy means for an organisation to gauge their own policy through self-assessment [15]. Unlike other measures, readability metrics can assist in improving the inherent properties of information security policies, 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 [9]. Ideally, institutions need to make certain 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 document and gauge how straightforward it is to comprehend the text. A readability index is evaluated by using statistical text analysis. Traditional 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 [16]. However, according to Gray and Leary [17], there are more than 220 factors that can affect readability. They classified the factors into four groups (Content, Style, Format, and Features of Organisation). An ideal readability metric would take into many variables into account for its measurement. However, in reality, this may be problematic if some factors prove intractable to easy measurement. Readability metrics usually return an approximation of a text's difficulty. This is often expressed as a grade level, i.e., the years of education study required to be capable of understanding the text [18].

E. Readability Metrics

There are many readability metrics but in this paper, we will describe two readability formulas: Flesch Reading Ease - a popular and widely used measure and the Strathclyde Readability Measure (SRM) - developed at our institution as a 'new generation' of readability metric. The SRM was selected for use in analysing our set of eight sample policies.

1) Flesch Reading Ease Formula (FRE)

This is one of the most commonly used readability formulas, 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 [19]. 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 [9]. Despite its popularity, in terms of common usage, the FRE has several recognised weaknesses in the selection of readability factors. A principal concern is the sole reliance on 'internal' characteristics of the considered texts.

2) Strathclyde Readability Measure

Of the available formulae for evaluating the readability of text, we adopted the Strathclyde Readability Measure (SRM) as it differs in approach from most other readability measures. Like the FRE, readability metrics often depend for measurement on counting syllables, characters per word, words, and sentences only, but SRM differs in taking into account the frequency of occurrence of words, relative to the British National Corpus, [20]. 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.).

F. Comprehension Test

Many test types that seek to measure understanding. The Cloze test is a method for gauging reading comprehension. Richards and Schmidt [21] explain that reading is the process of perceiving a text for the purpose of understanding its content. There is a possibility of achieving this silently, which is called silent reading. The understanding that results are 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), where the test creator decides which words are deleted based on some rational principle. The second technique is known as fixed ratio deletion or n^{th} word deletion, where every n^{th} word is removed systematically (at specific intervals). Thus, the test taker is required to construct meaning from the passage by identifying the missing words [21], [22].

III. RESEARCH DESIGN

This research was conducted in six broad phases to achieve the study objectives. In the first phase, 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 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 of similar word count (no more than 10% difference). In addition, university policies were chosen from top universities of 2015 - according to Quacquarelli Symonds (QS).

In the second phase of this work, without use of pre-determined inquiries, the 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 [23]. 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 information security policy 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.

In Phase 3, 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 [24]. In addition, focus groups can be used in combination with another method to clarify and evaluate research findings [25]. For these reasons, focus groups were used to validate the views expressed by the computer experts in the previous stage. Thereby, we have adopted a mixed method approach 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 [26].

Eleven people (professional computer users, who have used computers for more than six years), who expressed an interest in the case study, were invited to participate in a 90 minute discussion. This approach similar to the focus group meetings described by [27], [28]. Following on from this constructive dialogue, the focus group interview 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 was the easiest policy to understand from the set, 4) highlighting the variation between policies with respect to the aspects they covered.

Phase 4 focused on developing comprehension tests (Cloze tests). Following [21], [29]–[33], these Cloze tests are key to determining the comprehensibility of policy components and underpin the pilot study.

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 [18], [20], [34]–[36]. 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 [37], 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 [18].

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. In addition, Kobayashi [22] emphasises that there is a high correlation between readability metric scores and comprehension test results. 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 the readability factor may influence the efficacy of ISPs.

In Phase 5, as a pilot study, a number of Cloze tests were evaluated prior to the full-scale study. As noted by Connelly [38], 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.

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 [39].

Phase 6, the final phase, compared the comprehension outcomes against readability metrics to identify the similarity and differences between the results of human and software metric (SRM1 scale).

IV. Pilot Study

A. 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 understanding performance. In other words, the principal aim of the pilot study is to tune the subsequent process as much as possible.

B. Participants

A total of 20 university students participated in the study. These were all international students, as they were the target group for the study, in a mix of undergraduates and postgraduates. Each participant was randomly assigned one of two tasks. These comprised four Cloze tests and six multiple-choice questions (see section D for procedure detail).

Responses from three respondents were later withdrawn because of missing answers. The remaining data of 17 subjects were entered into the Statistical Package for Social Sciences (SPSS) program for analysis.

C. Materials

Reading passages were selected from the set of information security policies (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.

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 experts 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. Due to the small number of available

experts, this salient point determination process was an additional stage in which the focus group validated feedback from the experts.

It is worth mentioning that, this research used 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 'nth word' approach could select function words or other less significant aspects. For this reason, the decision was taken to avoid deleting proper nouns and numbers. An example of rational Cloze from one of the chosen policies is shown below.

Unlawful file-sharing using the University's information resources is a _____ of copyrights and _____ policy.

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.

In response to the respondents' feedback from the preliminary pilot study, the number of items for the Cloze test was set at 10, one item per ISP's statement. Overall, the reading material contained 80 items (10 blanks for each of 8 ISPs). After completion by the test subjects, all of the filled gaps were analysed and examined.

D. 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.

In the Cloze test pilot study, all eight of the chosen information security policies were used. The policies were anonymised to prevent any biasing influences on the participant's responses. Because it was impractical for each respondent to take all eight Cloze tests, the eight ISPs were divided 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. In spite of the fact that there was no time period set for survey completion, the members were urged to record their time. Once all of the participants had finished, all of the responses was marked and analysed by the survey designer.

E. Data Collection

The online survey was implemented and deployed utilising a specialised tool for creating online Cloze tests. The LearnClick software tool was selected for the task [40]. This features all question types and supports the creation of Drag and Drop Cloze items that are easily used by potential participants. Furthermore, LearnClick is able to create a test and store it in one place as well as the ability to store a large number of responses. The answers can be easily transferred into most common statistical formats. 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.

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). In fact, the target group of the questionnaire (international students) was specifically sought in the group selection process. 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.

F. 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 [41].

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

V. RESULT

A. Pilot Study

1) Demographics

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 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).

2) Cloze Tests

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.

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. 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 Ph.D., 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 I, with the values converted into percentages to make the comparison easier. Table II 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.

TABLE I. PARTICIPANT RESULTS

	Policy A	**Policy B**	**Policy C**	**Policy D**
Participants' count	9	9	9	9
Mean	6.1111	6.2222	5.5556	4.3333
Mode	7.00	8.00	8.00	4.00
Std. Deviation	1.26930	1.98606	2.69774	2.39792
Variance	1.1611	3.944	7.278	5.750
	Policy E	**Policy F**	**Policy G**	**Policy H**
Participants' count	8	8	8	8
Mean	7.3750	7.1250	7.0000	5.6250
Mode	8.00	5, 9 & 8*	6, 7 & 8*	6.00
Std. Deviation	2.44584	1.64208	1.77281	1.59799
Variance	5.982	2.696	3.143	2.554

TABLE II. DESCRIPTIVE STATISTICS

	Participants' count	**Mean***
Group		
First group	9	55.5556
Second group	8	67.8125
Qualification		
Ph.D.	2	65.0000
MSc	12	61.8750
BSc	3	30.0000
Computer Experience		
Proficient	7	64.2857
Intermediate	8	66.5625
Basic	2	30.0000
*Converted into percentages		

The results also show that participants' performance in Policy D were 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.

B. Readability Metrics

For our experiment, we used the Strathclyde Readability Measure (SRM1), as it is intended for text samples of more than 150 words. The following equation [20] gives the SRM1 score:

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

Where:
AWF = the average word frequency, only calculating words with a frequency not more than 100,000.

K = a constant depends on the average sentence length (ASL)

- 15: if the ASL is larger than or equal to 17 and less than 25, or the ASL is under 17 and the AWF is larger than 95000.
- 13: if the ASL < 17 or > = 25.

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 text. Generally, the range of Strathclyde Readability Measure differs from the Flesch Reading Ease formula in estimating readability (refer to Table III).

To obtain results for the eight policies, SRM software was used. The overall results, shown in Table IV, 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 E, Policy G, and Policy H. The SRM tool did not show any significant distance from the humans rating of the documents. Thereby, we can confirm that there are some correlations between the software readability formulas' results and human comprehension test results, and this supports our view that readability has an influence on understanding ISPs.

TABLE III. ESTIMATE OF READABILITY ON THE FRE AND SRM SCALE (SEE [9], [20])

FRE Scale		**SRM Scale**	
Mark	***Readability Category***	***Mark***	***Readability Category***
0-20	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 IV. COMPARISON OF HUMAN AND SRM1 RESULTS

Text	**Human**		**SRM 1**	
	***Mean*[1]**	***Rank*[2]**	***Rank*[3]**	***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

[1] Converted into percentages.
[2] One is the hardest text based on human prospective.
[3] 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.

VI. CONCLUSION

User compliance with information security policies 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.

This study adopted a number of methodologies to explore whether specific factors should be considered before releasing policies. 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 not giving sufficient 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, 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.

An extended version of this experimental approach is planned for the near future. This will cover wider population groups and look in detail at the correlation between survey variables. In due course, we aim to provide a set of guidelines that will permit checks on the likely readability of information security policies and thereby, allow for improvements in comprehension.

References

[1] S. Furnell, "IFIP workshop – Information security culture," *Comput. Secur.*, vol. 26, no. 1, p. 35, Feb. 2007.

[2] E. C. Johnson, "Security awareness: Switch to a better programme," *Network Security*, vol. 2006, no. 2. pp. 15–18, 2006.

[3] N. H. Higgins, "Corporate system security: towards an integrated management approach," *Inf. Manag. Comput. Secur.*, vol. 7, no. 5, pp. 217–222, 1999.

[4] G. H. Mc Laughlin, "SMOG Grading – A New Readability Formula," *J. Read.*, vol. 12, no. 8, pp. 639–646, 1969.

[5] A. ALArifi, H. Tootell, and P. Hyland, "A study of information security awareness and practices in Saudi Arabia," 2012, pp. 6–12.

[6] H. Chan and S. Mubarak, "Significance of Information Security Awareness in the Higher Education Sector," *Int. J. Comput. Appl.*, vol. 60, no. 10, pp. 23–31, 2012.

[7] B. Lebek, J. Uffen, M. H. Breitner, M. Neumann, and B. Hohler, "Employees' information security awareness and behavior: A literature review," 2013, pp. 2978–2987.

[8] R. S. Shaw, C. C. Chen, A. L. Harris, and H.-J. Huang, "The impact of information richness on information security awareness training effectiveness," *Comput. Educ.*, vol. 52, no. 1, pp. 92–100, Jan. 2009.

[9] F. Ammann and A. Sowa, "Readability as Lever for Employees ' Compliance With Information Security Policies," *ISACA*, vol. 4, pp. 1–4, 2013.

[10] K. Höne and J. H. . Eloff, "What Makes an Effective Information Security Policy?," *Netw. Secur.*, vol. 2002, no. 6, pp. 14–16, 2002.

[11] M. a. Alnatheer, "Information Security Culture Critical Success Factors," in *2015 12th International Conference on Information Technology - New Generations*, 2015, pp. 731–735.

[12] M. Chapple, "Four ways to measure security success," *TechTarget*, 2005. .

[13] ENISA, "Current practice and the measurement of success," *Eur. Netw. Inf. Secur. Agency*, no. July, p. 20, 2007.

[14] S. O'Bryan, R. Caraway, and C. CISA, "Critical Elements of Information Security Program Success," vol. 4, no. 26, p. 22, 2006.

[15] D. a. Chapin and S. Akridge, "How Can Security Be Measured?," *Inf. Syst. Control J.*, vol. 2, pp. 43–47, 2005.

[16] G. Weir and N. Anagnostou, "Collocation frequency as a readability factor," *Proc. 13th Conf. Pan Pacific Assoc. Appl. Linguist.*, 2008.

[17] W. S. Gray and B. Leary, *What makes a book readable*. Chicago: University of Chicago Press, 1935.

[18] N. Anagnostou and G. Weir, "From corpus-based collocation frequencies to readability measure," pp. 1–14, 2006.

[19] W. DuBay, *The principles of readability*. 2004.

[20] G. Weir and C. Ritchie, "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, 01-Jun-2006.

[21] J. C. Richards and R. Schmidt, *Longman Dictionary of Language Teaching and Applied Linguistics*. Routledge, 2015.

[22] M. Kobayashi, *Hitting the Mark: How Can Text Organisation and Response Format Affect Reading Test Performance? (Vol. 13)*. Peter Lang, 2009.

[23] D. Crosby, "Methodology for Eliciting Expert Opinion," *Mrc.ac.uk*, 2016. [Online]. Available: https://www.mrc.ac.uk/funding/how-we-fund-research/highlight-notices/methodology-for-eliciting-expert-opinion/. [Accessed: 03-Mar-2016].

[24] P. Liamputtong, *Focus Group Methodology : Principle and Practice*. SAGE Publications, 2011.

[25] Evaluation Research Team, "Data Collection Methods for Program

Evaluation: Focus Groups," *Eval. Briefs*, no. 13, p. 2, 2008.

[26] R. A. Krueger and M. A. Casey, *Focus Groups: A Practical Guide for Applied Research*. SAGE Publications, 2014.

[27] J. Bray, N. Johns, and D. Kilburn, "An Exploratory Study into the Factors Impeding Ethical Consumption," *J. Bus. Ethics*, vol. 98, no. 4, pp. 597–608, 2011.

[28] Z. Yunos, R. S. A. Hamid, and M. Ahmad, "Development of a cyber security awareness strategy using focus group discussion," in *2016 SAI Computing Conference (SAI)*, 2016, pp. 1063–1067.

[29] J. R. Bormuth, "Comparable cloze and multiple-choice comprehension test scores," *J. Read.*, vol. 10, no. 5, pp. 291–299, 1967.

[30] R. A. Guillemette, "The Cloze Procedure: Assessing The Understandability Of An IEEE Standard," *IEEE Trans. Prof. Commun.*, vol. 32, no. 1, pp. 41–47, 1989.

[31] M. Kobayashi, "Cloze Tests Revisited: Exploring Item Characteristics with Special Attention to Scoring Methods," *Mod. Lang. J.*, vol. 86, no. 4, pp. 571–586, Dec. 2012.

[32] E. F. Rankin and J. W. Culhane, "Comparable cloze and multiple-choice comprehension test scores," *J. Read.*, vol. 13, no. 3, pp. 193–198, 1969.

[33] W. Taylor, "'Cloze procedure': A new tool for measuring readability," *Journal. Q.*, vol. 30, pp. 415–433, 1953.

[34] G. R. Klare, *How to write readable English*. Hutchinson, 1985.

[35] S. A. Crossley, S. Skalicky, M. Dascalu, D. S. McNamara, and K. Kyle, "Predicting Text Comprehension , Processing , and Familiarity in Adult Readers : New Approaches to Readability Formulas," *Discourse Process.*, vol. 0, no. 0, pp. 1–20, 2017.

[36] W. DuBay, *The principles of readability. 2004*. 2008.

[37] G. Campbell and G. Weir, "Matching readers to texts with the Strathclyde readability measure," in *ICT in the Analysis, Teaching and Learning of Languages, Preprints of the ICTATLL Workshop*, 2006, pp. 49–55.

[38] L. M. Connelly, "Pilot studies," *MedSurg Nurs.*, vol. 17, no. 6, pp. 411–413, 2008.

[39] Z. Abu, H. Fracgp, P. S. Mmed, D. M. Fracgp, Z. Abu Hassan, P. Schattner, D. Mazza, K. Keluarga, and K. Lumpur, "Doing A Pilot Study : Why Is It Essential?," *Malaysian Fam. Physician*, vol. 1, no. 2, pp. 70–73, 2006.

[40] LearnClick Blog, "LearnClick Blog," 2016. [Online]. Available: https://www.learnclick.com/blog/. [Accessed: 30-Oct-2016].

[41] S. R. Terrell, "Writing a proposal for your dissertation: guidelines and examples," Guilford Press, 2016, p. 282.