2008

Case analysis of information security risk perceptions

Alexis Guillot
Edith Cowan University

Case Analysis of Information Security Risk Perceptions

By

Alexis Guillot

A Thesis submitted in partial fulfilment of the requirements

for the degree of

Bachelor of Information Technology Honours

(Computer Security - Internetworking)

Supervisors: William Hutchinson (Proposal) & Sue Kennedy (Thesis)

School of Computing and Information Science

Faculty of Computing, Health and Science

Edith Cowan University

USE OF THESIS

Copyright and Access Declaration

I certify that this thesis does not, to the best of my knowledge and belief:

(i) incorporate without acknowledgement any material previously submitted for a degree or diploma in any institution of higher education;

(ii) contain any material previously published or written by another person except where due reference is made in the text; or

(iii) contain any defamatory material.

Signed......[redacted].

Date.............28/08/2008.........

Abstract

The scientific rationality used by experts towards risk evaluation is expressed as the product of its likelihood of occurrence with its consequences or impacts (ENISA, 2006a). This directly opposes the subjective nature of risk perception, often appearing as inconsistent if not completely irrational (Byrne, 2003).

Risk perception theories are a pathway to explain the subjective nature of risk and a deeper insight into the human's cognitive system. Those theories may help to explain why people see, act and plan for risks in the way that they do, the weaknesses that exist in the human decision mechanisms and their impact on risk perceptions and decisions.

By questioning the existence of risk perception in the information security field of study, this research acknowledges those risk perception theories and provides a measure of their influence when rating information security risks.

The research measures the existence of risk perception issues by asking a participating sample of people to rate the likelihood of ten information security risks that carried previously measured statistics. In order to archive this, an online survey was designed to capture risk-rating information from an informed sample as well as a measure of their self-assessed information security knowledge.

By measuring the gaps between the participants' answers and the known occurrences of those risks, the research highlighted a number of disparities revealing the existence of risk perception divergence. A statistical analysis of the results was performed with the intent of highlighting gaps in the perception of the given risks. This analysis also allowed the research to narrow down the scope of risks that may or may not have been perceived with higher or lower gaps than other risks. Further analysis specifically identified the risks affected by those gaps, their statistical significance, strength and direction. The areas displaying the highest perception gaps resided with risks that were generally rare, new and unfamiliar or were being publicised in the popular media.

Finally, this research investigated whether or not the self-assessed respondents' knowledge is a factor influencing people's risk ratings in the online survey and thereby a factor in the way those risks are perceived. A correlation analysis was used to determine the degree of association between the participants' risk ratings and their perceived and self-rated information security knowledge.

Table of contents

List of Tables

List of Figures

1. Introduction

Risk assessment is a subset process of risk management and is extremely valuable to identify, analyse and evaluate risks. However, because of the general tendency for people to rely on such analysis and accept it as objective that it can be misleading. Furthermore, extensive psychological research has proved that risk analysis and assessment is very subjective, leading to the emergence and conceptualisation of biases, heuristics and risk judgments as important factors in risk analysis related disciplines.

Risk perception acknowledges that people judge risks subjectively based on their characteristics and severity, leading them to overestimate certain risks, and underestimate others. The inability of human perception to judge the reality of those risks has lead to the development of two major families of theory; the *psychometric paradigm*, concerned with heuristics, biases, cognitive and affective theories and the *cultural theory* **that** focuses on cultural rather than individual psychology. Both theories aim to explain why people assess the severity of risks differently. Therefore, to understand and consider the subjectivity and the existence of risk perceptions, as explained through those theories, should be highly beneficial to a risk analysis process.

By measuring the statistical gaps between the answers obtained from a participating sample to an online survey and professional surveys as sources of data, the research highlighted the existence of perception gaps in various risk areas. Risk perception theories reviewed and discussed in this research acknowledge that people judge risks subjectively. However, only broad risks are discussed by the literature. This research reconciled those theories and specialised risks by specifically focussing on information security risks.

1.1. Background to the study

According to the European Network and Information Security Agency (ENISA, 2006a) the risk analysis process involves:

- Examination of the risk sources;
- Their positive and negative consequences;
- The likelihood that those consequences may occur;
- The factors affecting those consequences;
- Assessment of the existing controls or processes that tend to minimize risks

While past experiences, market research, expert advice and security engineering models can provide a closer estimate of risk impact and its likelihood, the researcher believes that it is also crucial to assess risks at an individual and collective level (ENISA, 2006a). By considering the risk perceptions occurring at individual and collective levels, security professionals and risk decision makers would encompass further the scope of the risks and in turn, feed specific advice back to the risk analysis process for a stronger and more specific analysis. It is a reasonable assumption that individuals' perceptions might differ from reality itself.

This problem has been measured by numerous researchers and has been used extensively in psychology, social sciences, business (Burke & Greenglass, 2004; Osman, Gutierrez, Barrios, Kopper, & Chiros, 1998; Tversky, 1974), (Lerner, Small, & Loewenstein, 2004); other attempts were directed at security in broader terms (Schneier, 2007) but a general lack of information security related research prevailed. Understanding how those perceptions are triggereds and affect people's views is a very important aspect to any risk analysis.

1.2. Significance of the study

Organisations and users are exposed continuously to an endless number of new or changing threats and vulnerabilities that may affect their operation or the fulfilment of their objectives (ENISA, 2006b). As previously discussed, the management of such concerns is handled through the risk assessment process. Alberts and Dorofee (2001) describe risk assessment as a several step and complex multi-step process necessary to an effective risk management process.

Identification, analysis and evaluation of the threats and vulnerabilities posed to assets are the only way to understand and measure the impact of the risk involved and hence, to decide upon the appropriate measures and controls to manage them (ENISA, 2006a). Furthermore, risk assessment can be estimated using statistical analysis and formulas combining impacts and likelihood (Whitman & Herbert, 2005).

Risk analysis is a subjective process varying in detail according to the risk, purpose of the analysis, required protection levels and scope of the assets (ENISA, 2006a). The analysis can be qualitative, semi-quantitative or quantitative; however, a consistent approach with criteria defined by the risk management strategy must be followed. To fulfil this numerous frameworks and methodologies have been developed to formalise the approach to risk assessment (Alberts & Dorofee, 2001; AS/NZS:4360:2004, 2004; Octave, 2003) and all emphasise the importance of a broad knowledge of the organization's business and security processes.

However, Ames (2007) argues that the widespread application of mathematics makes quantitative risk assessment measurements appear more precise and mature than in fact they are, as they are often impractical and not cost-effective. In contrast, qualitative measures can be developed to be more granular and sophisticated; "the attitudes, perceptions, habits, history … will determine what approach will work" (Ames, 2007). This directly implies the integration of the human factor, as it is sometimes referred to in the risk assessment equation. After all, humans are often regarded as the weakest link in the security chain (Tan, 2006) and as poor judges of risks (Bailey, 2006). Schneier (2007), whose research outlines similar conclusions, argued that people do not evaluate security trade-offs mathematically or by examining the relative probabilities of different events; "instead, we have shortcuts, rules of thumb, stereotypes and biases generally known as 'heuristics'. These heuristics affect how we think about risks, how we evaluate the probability of future events, how we consider costs and how we make trade-offs". Schneier (2007) insists that considering the human factor, and more precisely the several ways one could perceive risks is of great importance therefore in order to understand all risks and thereby get the risks right.

Getting those risks right is important at a business level where an operational balance needs to be found between business processes and I.T. security. Considering the above can help us achieve this and enable the decision makers to take more informed security decisions. Similarly, security professionals can leverage the findings of the research in order to improve the risk assessment processes thereby to manage the risks better and to decide on the appropriate actions and levels of controls to put in place.

The intent of this research is to provide further information to help risk assessments by highlighting the existence of risk perceptions in information security. First, by analysing the data collected by a questionnaire, the research estimates the importance that risk perception may have on the representations of those risks. Then, through the analysis of those results, the research discusses factors that may influence or explain those perceptions. The objective was to determine the presence and impact of variables such as biases, heuristics and experiences that get in the way of risk assessment. It is from the results and concluding analysis of the findings that this research narrowed down those factors.

1.3. Research questions

Is there a gap between people's perception of information security risks and the findings of published research of those risks?

From this research question, the study answered the following sub-questions:

- Which risks are perceived with the highest, lowest, or no gap amongst all participants?
- Are there risk perception differences between people that self-assessed their information security knowledge 'above average' and the overall participating sample?

2. Review of the literature

2.1. *Information security surveys, source of risk measurements*

Measuring risk can be done in two ways; qualitatively, as outlined by the Australian Risk Management Standards (AS/NZS:4360:2004, 2004) using a model that rates risks as High, Medium, Low and quantitatively, using probabilities and statistics to mathematically measure losses, impacts and risks. The United States General Accounting Office explains that "efforts to develop precisely quantified risks are not cost-effective"; the amount of time and effort involved in the process are simply too great. (GAO, 1998). Qualitative measures on the other hand can be developed as a more granular and sophisticated approach. This research adopted a 'traffic light representation to illustrate the results of the data analysis (Section 5) in order to visually characterize the data. This method of representation is easily understood, visual and requires little detailed analysis (Ames, 2007).

Three criteria are noticeable when considering sources outlining risk measurements:

- People have different notions of what constitutes risk, threat and vulnerability (Ames, 2007);
- The surveys are an indicative measurement of security as it happens in a business context (Deloitte, 2006);
- "So far no one has come up with a common, clear, straightforward way of measuring risk that applies or works equally well in all circumstances" (Ames, 2007).
- Whilst the definition of risk, threat and vulnerability may vary from one individual to another, the questions used in this study resulted from risks specifically selected for presenting similar occurrences, shown by exhibiting similar percentages and ratings across all surveys.

Ames (2007) also argues that the science of risk became quite a separate discipline and highly mathematical, as a result many financial and economic components have been injected into risk components. However, the research does require a comparative basis that is:

- "Comparable" – which can be compared the scale of one risk meaningfully with another (Ames, 2007);
- "Consistent" – various surveys presented similar measures for the same risks (Ames, 2007);

- "Quantifiable" (Ames, 2007).

This is of a critical importance for this study. The interest directly lies in the risk measurements undertaken by those surveys. However, some surveys attracted criticism for unclear methodologies (Walsh, 2006); a preliminary research took special care to only retain risks that showed similar results statistics across all surveys. Furthermore, the AusCERT (2005; 2006) survey, as discussed in Appendix B, was adapted from the CSI/FBI Computer Crime and Security Survey (2006; 2007) which ensured high similarities between their methodologies, whilst providing the research with the assurance that their definition of the risks was consistent. The research focuses on risks that presented similar occurrence rates across both surveys as shown in Table 2-1.

Table 2-1 Risk occurrence percentages as per information security surveys

Selected risks (As per Appendix A)	AusCERT 2005	AusCERT 2006	CSI/FBI 2006	CSI/FBI 2007
Virus attacks	64%	66%	65%	52%
Self-propagating malware infection (virus or worm)	- %	45%	- %	- %
Non-propagating malware infection (Trojan or rootkit)	- %	21%	- %	- %
Denial of Service attacks	24%	18%	25%	25%
Theft or breach of proprietary or confidential information	14%	14%	9%	8%
System penetration by outsider	6%	7%	15%	13%
Laptop theft	53%	58%	47%	50%
Insider abuse of Internet access, e-mail or computer resources	68%	62%	42%	59%
Unauthorized access to information by insider	9%	8%	32%	25%
Website defacement	8%	8%	6%	10%

2.2. Conventional wisdom about risk

Nobel Prize winner Daniel Kahneman (2003) refers to the human brain as having two separate cognitive systems, one for intuition and one for reason:

> The operations of System 1 are typically fast, automatic, effortless, associative, implicit ... and often emotionally charged; they are also governed by habit and therefore difficult to control or modify. The operations of System 2 are slower, serial, effortful, more likely to be consciously monitored and deliberately controlled; they are also relatively flexible and potentially rule governed.

Bruce Schneier sees a causal relationship between those findings and the irrational trade-offs made by people. "Most of the time, when the perception of security doesn't match the reality of security, it's because the perception of the risk doesn't match the reality of the risk" (Schneier, 2007). To illustrate his arguments, he explains that food poisoning kills an average of 5000 people every year, and automobiles killed another 40,000; yet the public seem to be more worried about the non-repeated 9/11 terrorist act which in comparison caused less than 3000 deaths (Schneier, 2007). Looking at those statistics, he comments that people are more afraid of flying than driving their cars. Those issues lead Schneier to construct Table 2-2

Table 2-2 Conventional risk wisdom

People exaggerate risks that are:	People downplay risks that are:
Spectacular	Pedestrian
Rare	Common
Beyond their control, or externally imposed	More under their control, or taken willingly
Talked about	Not discussed
Intentional or man-made	Natural
Immediate	Long-term or diffuse
Sudden	Evolving slowly over time
Affecting them personally	Affecting others
New and unfamiliar	Familiar
Uncertain	Well understood
Directed against their children	Directed towards themselves

Morally offensive	Morally desirable
Entirely without redeeming features	Associated with some ancillary benefit
Not like their current situation	Like their current situation

2.3. Risk perceptions theories

Risk perception is subject to judgment that people make about the characteristics and severity of a risk. Numerous theories emerged to explain why people make different estimates of the dangerousness of risks, however social scientists have developed two major families of theory, namely the psychometric paradigm and cultural theory (Slovic, 1987; Thompson, Ellis, & Wildavsky, 1990).

The psychometric paradigm has been developed "in order to explain risk evaluations in terms of intuitive mental rules-of-thumb that people use when judging the likelihood of everyday events" (Skjong & Wentworth, 2001). It mainly focuses on three areas, the heuristics and biases (Kahneman & Tversky, 1984) the cognitive theories (Slovic, Fischhoff, Lichtenstein, & Roe, 1981) and the affective theories (Finucane, Slovic, Mertz, Flynn, & Satterfield, 2000). The cultural theory on the other hand is concerned with cultural biases (Douglas, 1994).

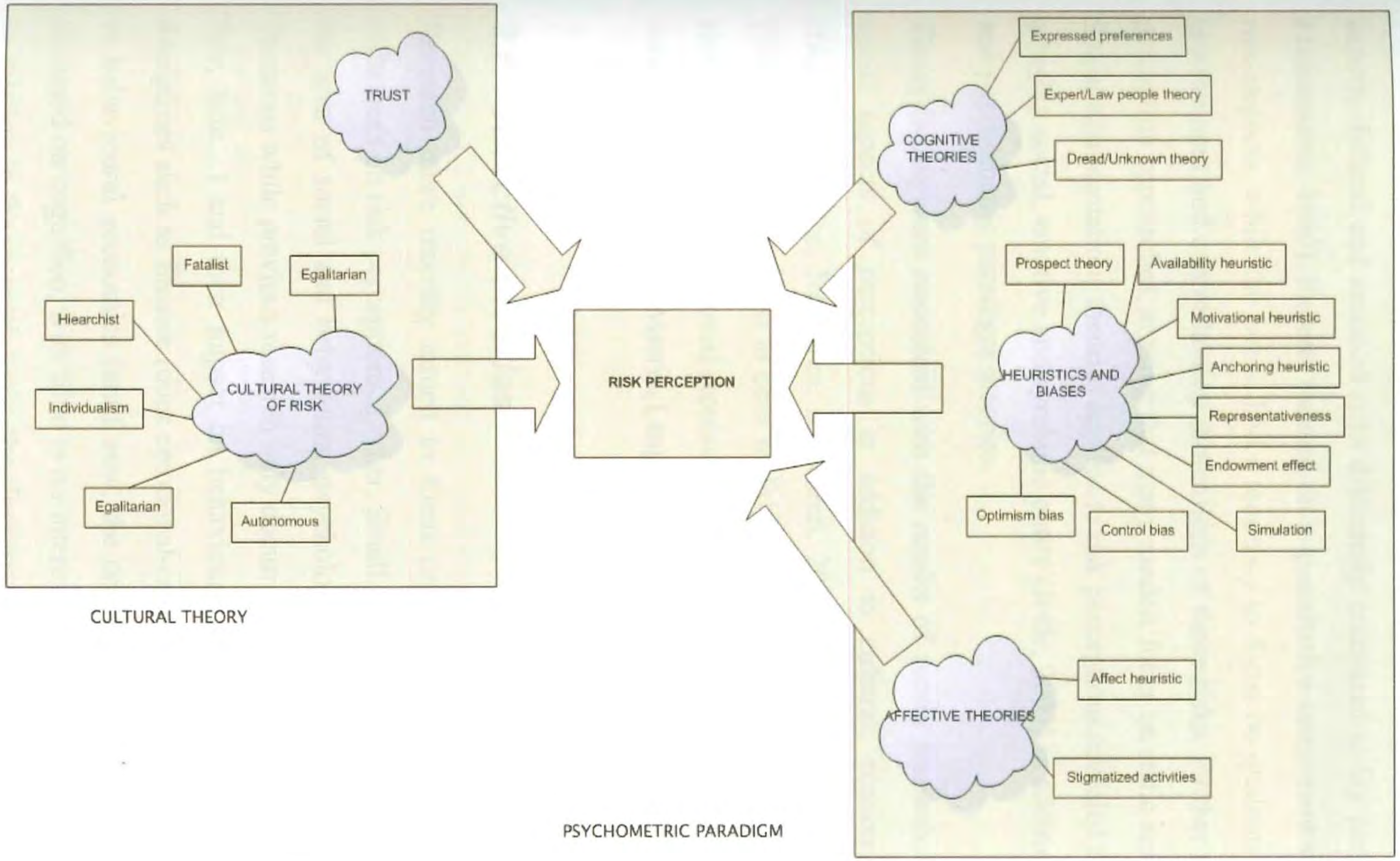

Figure 2-1 Risk perception theories: This figure summarises the complex and multi-layer approach existing within the two major risk perception concepts. Both Psychometric paradigm and Cultural theory are divided into sub-researches, here represented by the 'clouds'. Each of those research domains saw the emergence of sub-theories or concepts, represented here as being attached to the 'clouds'. Some of those concepts are discussed through the literature review of this research.

2.4. Social perceptions and representations

In addition to the emotional, economic and various factors affecting risk perceptions, social psychologists have been interested in the social aspects of perceptions. Research found that experts defined and assessed risks differently compared to lay people (Slovic, Fischhoff, & Lichtenstein, 1980). Experts focused on a quantitative assessment such as risk likelihood and risk impacts, while the public had a tendency to focus on qualitative arguments, namely the involvement and controllability dimensions of those risks. Further research explored the risk perception approaches, arguing that many models focus on static and intra-personal processes. Social representations theories applied to risk perceptions revealed that risk responses are also a highly social, emotive and symbolic entity (Joffe, 2003) in addition to the common cultural and psychometric paradigm theories.

Those findings are consistent with the results of recent research, which all considered the social aspects of perceptions in addition to cultural, economical and emotive issues (Dijksterhuis, Bos, Nordgren, & Baaren, 2006; Hasehuhn & Mellers, 2005; Lerner et al., 2004) hence, the interest in both social representations and risk perceptions. Also it was felt strongly that bridging social representation theories to risk perception would bring to it a more sociological approach (Marris, Langford, & O'Riordan, 1998).

2.5. Affective theories

Research more recently turned to focus on the effects of emotions and their roles and influences on risk perceptions. Lerner, Small, and Loewenstein (2004) created a precedent in the area of social and behavioural psychology by linking emotional impacts to economic decisions while previous research only documented the carryover of specific emotions (anger, fear, hate...) and their impact on behaviours. The findings have ramifications in various disciplines such as finance (does emotional-carryover diminish when real money is at stake?) or behavioural economics (until now, the decision making processes have been essentially focussed on cognition, with little to no interest for the emotions involved in those decisions). In relation to the research topic, the findings are useful as Lerner, et al provide this research with a basis to suggest that emotions can have a dramatic effect on transactions even though they arise from a prior, irrelevant situation. This is a direct implication of the emotional and social theory.

Other studies (Dijksterhuis, Bos, Nordgren & van Baaren, 2006) have used Lerner, Small and Loewenstein's results as a starting point for further investigation, providing additional layers of acceptation and peer-reviewing to the study. The initial findings served as a great complementary basis to Dijksterhuis' "deliberation-without-attention" effect, which introduces the impacts of conscious and unconscious deliberation on decisions. The latest dictates that complex choices were viewed more favourably when decisions were made in absence of attentive deliberation, which are partly explained by the emotional states of the tested subjects as seen in Lerner, et al (2004).

The work of Lerner, et al. found a noticeable extension in the research lead by Haselhuhn (Hasehuhn & Mellers, 2005) which complement their conclusions on the emotional implications of their subjects into economical activities. Those theories also find many correlations to major works on risk perceptions, risk judgments and influential factors in decision-making (Kahneman & Tversky, 1984; Slovic, 1987; Tversky, 1974).

2.6. Theoretical framework

Figure 2-1 established that risk perceptions are subject to a variety of factors of influence.

To complicate the situation, often more than one factor often influences a participant in a given situation. The multiple influential factors were also found to be subjective, depending on the audience sampled (Slovic et al., 1980). A well as highly dependent on the risk presented to the participants, factors such as risk magnitude, impacts, economic or emotional involvements all played important roles (Dijksterhuis et al., 2006; Hasehuhn & Mellers, 2005; Lerner et al., 2004; Tversky, 1974).

This study is interested specifically in information security risks. The research aim is to highlight the existence of perceptions directed towards those risks. The following research framework is designed exclusively to provide a specific and dedicated focus on each of the ten analysed risks.

The information security surveys from AusCERT 2005, 2006 and CSI/FBI 2006, 2007 were used as sources for risk quantification. The AusCERT and CSI/FBI surveys questioned the participants on the nature and type of attacks detected over a 12 month period. The contributors' answers, formulated into percentages, outline the occurrences of the attacks. Those percentages provide the comparative basis for this research to establish the "riskiness" of the threat driving the attacks. As previously discussed, the surveys acted as a reference as well as a comparative baseline to highlight the gap between the participants' answers and the results of those surveys.

The following framework was designed to answer the research question; it applied to each one of the ten risks. Once data for all ten risks were registered, the research used the same workflow and applied it to each respondent.

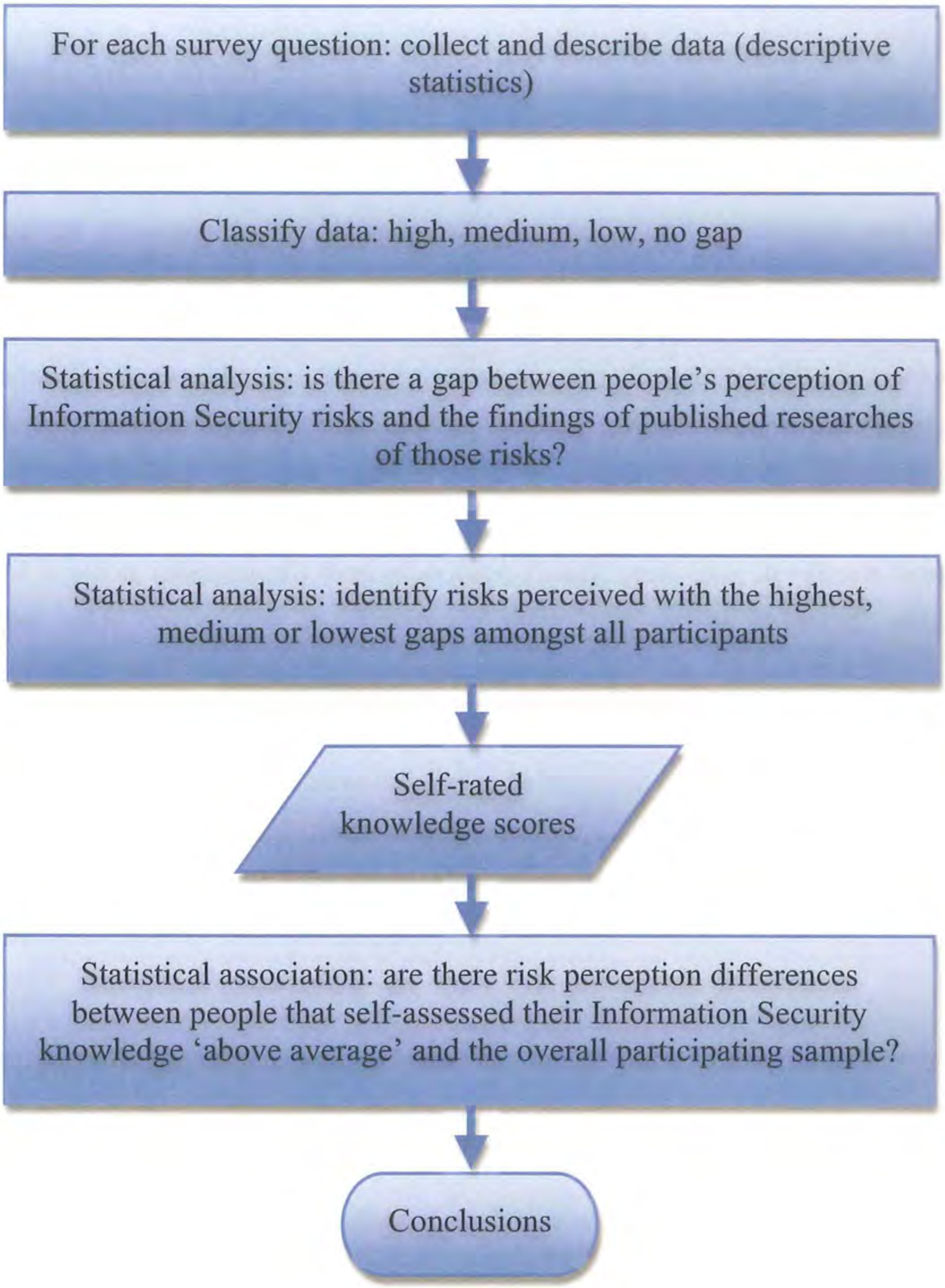
For each survey question: collect and describe data (descriptive statistics)
Classify data: high, medium, low, no gap
Statistical analysis: is there a gap between people's perception of Information Security risks and the findings of published researches of those risks?
Statistical analysis: identify risks perceived with the highest, medium or lowest gaps amongst all participants
Self-rated knowledge scores
Statistical association: are there risk perception differences between people that self-assessed their Information Security knowledge 'above average' and the overall participating sample?
Conclusions

Figure 2-2 High-level research framework: This figure is a high level overview of the research and how the several elements of this research, starting with the data collection to the conclusions, interacts and flow.

3. Research methods and design

3.1. Research design

The branch of philosophy known as epistemology is dedicated to knowledge and more precisely, to how we come to know. Methodology, on the other hand, whilst also being concerned with knowledge is much more practical and focused on the specific ways one could use to understand the world better. These two concepts are intimately related, one involving the "philosophy of how to know the world while the latter involves the practice" (Trochim, 2006c). In epistemology one view, referred to as empiricism, argues that knowledge is derived from experience; whereas rationalism considers knowledge as acquired through reason and intuition, independently from experience. Finally, positivism, an extreme form of empiricism, holds that nothing is inherent and things that can be measured, such as observation and experimentation, only acquire knowledge.

In order to comply with this, this research used a positivist epistemology as the knowledge required to address the research question requires to be measured through experimentation. In addition, the quasi-experimental design was selected to allow the researcher to present arguments from the literature to infer causal relationships whilst opening the research design to the opportunity of further data collection.

3.2. Research methodology

For this research, the quantitative data collection and analysis is based on the administration of a survey instrument.

The use of a survey instrument as a data collection means requires a level of interaction between the researcher and the participants. Trochim's methodology (2006d) was used to determine the appropriate classification of this research. In accordance with his method, the research design does not use random assignment to groups; however, it does use multiple groups or multiple waves of measurements. Therefore, the research methodology that is best suited to this research is the quasi-experimental design. This allows the researcher to have samples that are not a true representation of the population. Furthermore, the lack of true randomness from the sample selection means that causal relationships can be inferred only, not used to prove causality.

The data gathered from the survey is designed to:

- Capture the self-rated knowledge in information security of the participating sample;
- Record the participant's own rating of given risks' instance; &
- Capture this rating within the context of the likelihood of the occurrence of those risks.

The analysis of the gap between the participant answers and the known results to the questions allowed the research to correlate this data between the participants, depending on the results, and infer causal relationships with the participant's self-assessed levels of knowledge in information security. The basis for the analysis of that data is outlined in Section 3.7.

3.3. Survey design

The quantitative nature of the data collection lead this research towards the adoption of a data collection process performed through a survey.

Surveys are a means of collecting data for descriptive purposes. For the purpose of this research. Closed-ended questions were used in order to provide the participants with a limited number of possible alternatives. The participants were asked to choose one answer only, the answer that best represented their beliefs.

This was controlled within the survey software as participants were able to select one answer per question only. The survey technical setup is further discussed in Section 4.1.

The survey construct made use of two types of rating scales discussed in Section 3.6.

The chosen survey administration method was an anonymous web questionnaire. The procedure for data collection and survey administrating can be found in Section 7.6, whilst the technicalities and configuration details of the survey software used to collect the data are available in Section 8.1.

3.4. Sampling methodology

The methodology employed allows the sample to not be a true representation of the population. This research uses a non-probability purposive sampling method (Trochim, 2006b). Such a sampling technique is used when the individual members of the population do not have an equal likelihood of being selected to be a member of the sample (Jackson, 2006). The validity and reliability of this sample is further discussed in Section 3.8 and 3.9.

The type of non-probability sampling selected for this research is known as quota sampling. It involves ensuring that the sample is similar to the population in certain characteristics. Even though this research tried to ensure similarity with the population in certain characteristics, the population is not sampled randomly. Participants are found wherever they can be, through whatever means is convenient (Jackson, 2006).

A selection of an informed sample fulfils the unique purpose as being computer literate is the only characteristics in which this research is interested. This direction was chosen to ensure similarities with the information security surveys used, although limited information was available on their original targeted samples. This issue is further discussed in Section 3.8, 3.9 and 3.10.

The target sample received a standardised e-mail invitation to participate in the survey. The invitation contained a link to the questionnaire as well as the content of the information letter that outlined a brief explanation of the study, the privacy conditions surrounding the data collection, as well as the consent request. That information is available in Appendix C. Participants were encouraged to provide honest and accurate answers as the data collection was anonymous. Consent was mandatory in order to begin the survey.

The targeted population for the survey component consisted of 48 individual participants (N=48) drawn from various sources. The invitations were sent to various security groups in the community that include security researchers, academics and professionals from public and private sector. Representatives from the School of Computer of Information Science Security Research Group (SCISSEC, 2007), which includes seven lecturers and a dozen research students at both undergraduate honours and postgraduate level agreed to notify the members of their groups and associations about the survey and encourage them to participate. Analogous associations, namely the West Australian Information Security Special Interest Group (WAISSIG, 2007) and the Australian Information Security Association (AISA, 2007) were amongst the represented groups. In addition to those security professionals, survey invitations were generated for an equal number of individuals that were known *not* to belong to security groups. The intent was to create a balance of the represented population with both security professionals (or those related to security group or activity) and participants that were not.

The use of a relatively informed and professional sample should ensure a high response rate. To maximise this rate further, the participants had a two weeks period to complete the survey. In addition, by hosting the questionnaire online, it was available 24/7, effectively allowing the participants to answer the questions at their convenience. The validity of the sample is discussed in Section 3.9.

3.5. Procedure for data collection

- The questionnaire was hosted on a secure server hosted internally by the School of Computer of Health and Science (SCIS) which has endorsed similar data collections in the past and has the infrastructure available to support the needs of this questionnaire;
- A secure URL linked the participants to the questionnaire in order to ensure the security and privacy of the data collected;
- A random number was generated for each invitation and aggregated into the secure URL in order to directly link participants to the questionnaire; this ensured that each participant only completed the questionnaire once;
- The participants' answers were collected in a password protected database;
- The credentials were available to the researcher only, providing additional privacy guarantees;
- Participants were identified only by a unique computed generated number when they input their records via the online survey;
- The survey was made unavailable and removed from its host after a two week period;
- Information was collected and transferred to removable medium where it was encrypted; &
- All collected data will be preserved for a period of not less than five years in order to comply with Ethics requirements as outlined by Appendix D.

Further information supporting the above section is available at the following locations:

- Appendix A: Questionnaire;
- Appendix B: Information security surveys selection criteria;

- Appendix C: Information letter and informed consent;
- Appendix D: Ethics approval letter; &
- Section 4.1: Questionnaire setup.

3.6. *Sampling measurements: rating scales*

Trochim (2006a) outlines the broad steps towards the development and usage of a rating scale with which this research as follows:

- Define the focus of the scale;
- Generate the items which define the percentage ranges used by the scale; &
- Rate the items that provide the measures required by the research.

This led to the creation of two scales: one for the measurement of the participant's knowledge in the information security discipline and one which focused on a qualitative risk rating in terms of a percentage range.

Ordinal scales, often referred to as *ranked data*, are being used for all eleven questions. Objects are assigned to categories that carry a numerical property and form a rank order along a continuum. The data have the properties to identify and rating the magnitude but lack equal unit size and absolute zero (Jackson, 2006).

Question 1, representing a self-assessment of the participants' knowledge was rated against the following nominal scale:

Table 3-1 Self-rated knowledge scale

No knowledge	0
Low knowledge	1
Some knowledge	2
Average knowledge	3
Good knowledge	4
Expert knowledge	5

This type of scale is known as an *itemized rating scale*, or *specific category scale* (Krech & Crutchfield, 1948) used in the quantification of judgments and social attitudes.

Those scales are used as the self-reported measure on how people report that they act, think, or feel (Jackson, 2006), thus aiding this research in collecting data on cognitive events by asking individuals to report how they rate their information security knowledge. Precautions were taken to ensure the reliability and validity of the self-ratings and are discussed in Section 3.8.

Questions 2 to 11, present a series of ten risks to the participants. From the survey's point of view the respondents were asked to rate them against the following five-point ordinal scale:

Table 3-2 Risk rating scale

Very low	0-20%	(light green)
Low	21-40%	(green)
Medium	41-60%	(yellow)
High	61-80%	(orange)
Very high	81-100%	(red)

The scale scores are derived from percentage range suitable to the survey findings as shown in Table 2-1. The reliability of those scales is discussed in Section 3.8.

3.7. Data analysis

From the data analysis point of view, the ordinal scales used for all eleven questions carried identity and magnitude for each question. Each individual item received a rank (i.e.: a number) that carried identity and conveyed information about order of magnitude (i.e.: how many participants rated themselves as having high, low, very low knowledge within the sample group).

Quantitative methods applied to the analysis of the collected data (discussed in Section 3.7).

Two parameters in the targeted population are at the core of this research:

1. The parameter of self-assessment rating (measured as an ordinal value, i.e.: a number)
2. The parameter of the likelihood of a risk in the opinion of the participant (measured as an ordinal value, i.e.: percentage rank).

Taken individually, the second parameter is the one that provides this research with the basis to partially answer the following research questions:

- Is there a gap between the overall participants' answers and the results of previous surveys to the questions on risks?
- Which risks are perceived with the highest, lowest, or same gap amongst all participants?

However it is the relationship between those two parameters that is the basis to the remaining research question:

- Are there any risk perception differences between people that self-assessed their information security knowledge as being 'above average' and the overall participating sample?'

Descriptive statistics were used to conduct the preliminary analysis for Question 1 to 11 and describe the results from the survey construct in meaningful terms. This analysis allowed the research to depict the data set in terms of numerical measures that describe the *distribution* and *central tendencies* for the answers to each of the eleven questions (Section 5.1).

Inferential statistics methods were used to draw observations from the data collected and answer the research questions.

It is critical to the analysis that the data collected from question 1 is not directly compared to the results collected for the remaining questions. Similarly, the research did not categorise the participants based on those data. Therefore, answers provided by question 1 were not used to judge the level of knowledge of the participants, as opposed to a self-assessment, against their performance on the remainder of the questions, nor did it play any role in categorizing the participants' knowledge levels based on their answers. The intent behind the self-assessment question is to provide an additional variable for the analysis and highlight of information security risk perception and thereby answer the research question: 'Are there any risk perception differences between people that self-assessed their information security knowledge as being 'above average' and the overall participating sample?'

Question 1 however was used as a comparative baseline to evaluate the *association* between two variables using *bivariate correlations*. As seen in Section 2.4, social perceptions and representations also play an important role in risk perceptions. It therefore represents a variable of interest that this research measured. In order to do this, a calculation of the *correlation coefficient* was used to determine the degree of association between two variables of interest, thereby determining their statistical independence and measure of association.

For the purpose of this research, the scope of the analysis was limited to the following two variables:

- Participant's self-rated knowledge score and
- Participant's answers to ten risk questions.

Questions 2 to 11 outlined a measure of the *gap* between an individual's perceptions of risks against their ranked occurrences as shown by the information security surveys.

The data for each question was organised and described using *descriptive statistics*. *Inferential statistics* were used to test the data for variance with the results extracted from the information security surveys (providing a measure of the *gap*).

In inference testing the significance of this *gap* is demonstrated and interpreted for questions 2 to 11 through the comparison of the reference baseline's mean ($\mu 0$), (as shown in Table 2-1), with the calculated mean for the survey sample ($\mu 1$), (performed in Section 5.2).

The *one-sample t-test* was one of the methods used for inferential testing. The *t-test* is a parametric inferential statistical test of the null hypothesis for a single sample where the population variance (or *gap*) is not known (Jackson, 2006).

The *one-sample t-test* allowed the research to measure the variations between the mean ($\mu 0$) obtained for the reference information security surveys (Table 2-1) and the mean ($\mu 1$) calculated for questions 2 to 11 (from the survey sample group) as seen in Section 9.2. This effectively provided a measure of the *variances* between the means, thereby revealing gaps in the perceptions of risks.

The *one-sample t-test* also provides this research with a measure of the *two-tailed hypothesis test (p)*. The significance of this *p* value is what further confirmed the *variance* or *gap* in risk perceptions.

The samples were later categorised into groups of different gap strength to show an easy and visual comparison. Each group was associated with an ordinal metric value defined by its *variance* with the reference answer from the surveys (shown in Table 2-1) as follows:

- High gaps – participant answers have a variance of +/- 3 points on the five-point-scale.
- Medium gaps – participant answers have a variance of +/- 2 points on the five-point-scale.
- Low gaps – participant answers have a variance of +/- 1 point on the five-point-scale.
- No gaps – participant answers are the same and have no variance

3.8. Reliability

"Reliability of a measure is defined as the extent to which it is free from random error components" (Judd, Smith, & Kidder, 1991).

The risk definition, as seen in Appendix A, may sound unitary or multi-dimensional to the participants. To ensure the reliability of the risk definitions, the risk formulations used for question 2 to 11 (Appendix A) are comparable with the risk definition and wordiness used by the information security surveys. Furthermore, as the data analysis phase compares each participant's answers to the known rankings in the professional information security surveys as seen in Table 2-1; the data analysis demands the use of identical risk wordings.

Shrauger and Osberg (1981) outline several precautions that must be taken to obtain reliable and valid self-ratings. Following their recommendations, in this research the individuals were explicitly told the attributes to be rated; for example in the case of question 1 "your information security knowledge". The second recommendation is concerned with the accuracy of the answers. It is often believed that self-raters distort their responses rather than convey their honest assessments if those ratings are to be used to distribute valued resources (Judd et al., 1991). The authors insist that evidence for this is limited. However, this research enlists the participants' co-operation in the rating task by stressing that the accuracy of the answers is highly valued and that the answers will remain completely confidential and anonymous. This information is a part of the consent page as shown in Appendix C.

To ensure the reliability of the scale used by questions 2 to 11, Judd, et al (1991) advise the construction of a scale with a wider range, as it may make the participants more comfortable in indicating their position and help reduce biases. The halo bias, according to Cooper (1981), which is the tendency for an overall positive or negative rating of an object, can become problematic and needs to be considered in the formulation of the survey construct. To name a few halo biases, the generosity error leads to an overestimate of desirable qualities of an object whilst the contrast error might lead the participants to see an object as opposite to them on a trait (Judd et al., 1991).

Finally, the motivation of the participants to provide honest input was found by Cooper (1981) to be beneficial in reducing those biases. This was previously discussed and addressed by the construct of the survey.

3.9. *Validity*

The validity is the extent to which a tool measures what it is intended to measure (Tucker, 2007). In the context of this research the validity of the findings is highly dependent on accurate perception measurements.

The ability to demonstrate close similarities between the survey measurements shown by Table 2-1 represents a type of validity test. Guillot and Kennedy (2007) showed that both AusCERT and CSI/FBI used an informed sample in the public and private sector with participants that held various targeted positions; this is also shown in Appendix B. This research selected a similar sample (as shown in Section 3.4) that identifies the participants as members of the security community, thereby providing the research with an informed group of participants who are familiar with security related risks. Furthermore, the survey provides the opportunity for the participants to answer the questions from several points of views, thereby setting a context for the given risk. The participants are able, if they choose to do so, to make a distinction between their perceptions of the given risk in relation to those contexts (risks as perceived to affect the participant at a personal level, at a professional level, and the risks as perceived to affect the world in general).

It appears that these measures help to ensure a high degree of reliability and validity of the construct and the overall research.

The AusCERT 2006 survey includes the risk occurrences from 2003 to 2006. CSI/FBI on the other hand limits the display of its results to the current year. Furthermore, the CSI/FBI 2006 and 2007 editions contain specific percentages of occurrence for each risk in addition to a graphical representation of all risks. However the 2005 edition of the survey was limited to the same graphical representation and did not contain a measure of exact percentages for each risk. Given that the researcher could only determine an approximate value of those percentages, the decision was made not to include the results from CSI/FBI 2005.

The ability to provide information on the risks over various years allowed the researcher to have comparative values that could indicate the evolution and trends of the risks over time. This is also a mitigating factor to halo bias and tendencies affecting reliability. By ensuring that participants' perception of risks is not assessed against one specific occurrence nor a specific year, rather over a few years and a few sources of risk measurement. This allows more flexibility and scalability in the research and analysis of the gap between the survey's information and the participants input.

Moreover, the specially designed five-point scale allows a degree of flexibility by having percentage ranges that were representative of all information security surveys. The research also has the ability to mitigate the limitations originating from the surveys in terms of accuracy of their risk representations by having a broad range of percentages, t. Those ranges, associated with their quantitative equivalent (very low, low, medium, high, very high) allow a more granular and sophisticated approach to risk representation for the analysis.

The validity of the data analysis relies heavily on the definition of risks. In order to mitigate this and allow a proper comparison of this research with the results, as per the information security surveys, the definitions of the risks in the ten risk questions (as well as the wordiness of the questions) must match their equivalents in the information security survey. Skjong and Wentworth (2001) stated that the word risk is known to be ambiguous and not always as specific as being the product of probability/likelihood and consequence/impact as typically defined by risk assessment.

The researcher acknowledges that the risks as defined in both surveys and questions 2 to 11 are subject to the interpretation of the participant. Therefore, the validity of the results is subject to each participant's understanding of the question.

Nonetheless, the results of those questions are analysed and compared to the results shown by the information security surveys, (summarised in Table 2-1). Therefore, integrity and coherence of this analysis demand the usage of the same risk statements.

In order to mitigate further potential construct validity issues arising from question misinterpretation, a context (the participant's personal opinion) for the answers is defined within the each of the questions. This effectively limits the scope of the interpretation that the participants might have of a given risk.

The context provides an additional validity layer as it helps the participants who may have different levels of information security knowledge, rendering the survey more accessible to a larger sample. Furthermore, by asking the participants' opinion directly the question hints for their perception of the risk (i.e.: "in your opinion").

This research aim is to measure the gaps between the respondents' answers to the questionnaire and the information drawn from the information security surveys. The broad nature of this scale enabled the research to reveal those gaps whilst opening the analysis to the presentation of a variety of factors and perception related concepts. Furthermore, the scale allows the research to be preserved from narrow conclusions that would deserve much more expert attention from psychology researchers. Additionally, this study focuses on a specific aspect of risk measurement: the participants' perceptions. Therefore, the findings can be interpreted only within the context of the sample used. Generalisations can only be relevant in the light of further research on larger samples.

The questionnaire was purposely designed to with the self-rating knowledge question, as the participants were likely to judge their capabilities in a more accurate fashion, if the scope of the remaining questions were left unknown.

It is also important to mention that the research disregarded any emotional aspect at an individual levels to enhance the validity of the analysis. The focus of the research remains the highlight and measure of gaps in risk perceptions.

The comparison of the measured perceptions and those outlined by the surveys is the preliminary step which could, in time and with further research, be extended to target the emotional aspects and involvements of those same information security risks. Emotions and risk perceptions are subjects covered by the literature (Lerner et al., 2004) as discussed in the literature review. Whilst those ideas may be used in an analysis of the information, this research was restricted in order to avoid considering any emotional aspects for individual respondents. Any comment made in relation to emotional involvements influencing information security risk perceptions were formulated as hypothesis, or in relation to existing researches, in the context of all participants only.

3.10. Limitations

This research only intends to be a preliminary quasi-study to highlight the existence of risk perceptions in information security.

Whilst the literature on the psychology of risk and risk perceptions is abundant, the researcher's background is in computer and information security, not in psychology. As per the research questions, scope and time constraints, the research do not cover the interpretation of the results by other means than a statistical analysis and description of the data. The researcher acknowledges that is a potential limitation of the study, however, the complex nature of this analysis demands that it takes place in the context of psychology and related fields of study. It would also require more attention than the scope and timeframe allocated to this research allows. However this research does provide the basis to highlight the existence of risk perception gaps and a narrower focus on the nature of the affected risks (Section 5.4).

This approach may facilitate the focus of further work on the highlighted risks and may effectively narrow down the scope of further analysis (i.e.: require more centred questions around the highlighted risks or begin a psychological analysis and interpretation of the existing risk perception gaps).

Another limitation is the focus on specific sources of information for the measurement of the risks outlined in the questionnaire. The methodologies of those surveys is often criticised and showed some limitations (Brenner, 2006; Walsh, 2006; Winkler, 2006). On the other hand, the surveys presented several attractive and beneficial aspects to this study such as a comprehensive presentation of risk occurrences both locally and internationally.

This research acknowledges that the results presented in this document are best interpreted in their intended context: considering the results from the information security surveys.

The self-rating of information security knowledge by the participants is subjective and therefore subject to biases from the participants themselves. Whilst preventive actions were taken to minimise those biases and ensure that the reliability and validity of the self-ratings (section 3.8, 3.9), it is also considered as a limitation to the research that must be acknowledged.

A limitation specifically applying to the statistical data analysis is the lack of precise data from the information security surveys to provide accurate inferential statistics and compare the results of the online survey to the hypothesised baseline (CSI/FBI and AusCERT).

A *one-sample t-test* is the method used to determine whether the mean of the population from which the sample was drawn is the same as the hypothesised mean. Section 5.3 present the *gaps* between the means on a per question (or per risk) basis and clearly shows that the mean calculated from the survey construct (Section 5.2) provide much more detailed information than the ones offered by the information security surveys. For instance, the means calculated using SPSS for Windows offer the granularity of several decimal places as opposed to the data extracted from the information security surveys that were only in percentage form. Those percentages were later converted into an interval scale for the purpose of the analysis (as explained and detailed through Section 3.6 and 5.2). This research had to operate with the assumption that the scales conversion process would not attract an important error margin and thwarts the statistical analysis. Had the information security surveys fully disclosed the results per respondent (as per Appendix E) this research would have been able to directly apply the *one-sample t-test* method to compare the mean score from the online survey to the mean score calculated from the information security surveys.

3.11. Research methods summary

Research design:

- Quasi-experimental;
- Quantitative survey (descriptive methods for survey construction);
- Quantitative data analysis (using descriptive and inferential statistics).

Type of survey method:

- Anonymous web questionnaire (e-mail invitations).

Sampling technique:

- Non- probability (quota sampling).

Question type:

- Self-reported or Itemized rating scale;
- Close-ended questions with rating scales (ordinal scales).

4. Data collection

The survey software used to conduct this research was the open source package known as "Unit Command climate assessment and Survey System" (UCCASS, 2004).

The UCCASS (pronounced [yoo-kas]) is a PHP based survey script that allows the creation of online surveys. It was originally designed for organizations to administer Command Climate Assessments (annual surveys), however, the program can be used to create any type of online survey or questionnaire.

A full range of options and access control restrictions are available from the web-interface, the ones with a direct impact on both ethical requirements and validity of this research are outlined below.

4.1. Survey setup

The UCCASS (2004) version used for this research was v.1.8.1.

Noticeable configuration changes made by the researcher included:

- The survey setup was specifically limited to "e-mail only".
- The invitation codes were made of a randomly generated string of ten (10) alphanumeric characters in order to maximise the entropy of each invitation code.
- No participants were allowed to view the overall results of the research, only a summary of their own answers.
- The setup option of allowing participants to take the survey only once (per-invitation) was enabled.
- It was mandatory that the participants comply with the choice of one answer only; selecting more than one answer per question would trigger the answers not to be recorded and display an error message.
- Furthermore, the survey software allowed the questionnaire to be hidden from the outside world, so that it was accessible only through the e-mail invitation URL. This option was enabled.

- A re-direction page [index1.php] was created to direct the participants, to a greeting page, thanking them for their participation at the end of the successful completion of the questions.
- The survey was automated to activate on March the 25th 2008 at 6.00am DST (25-03-2008 @ 6.00 DST) and send the e-mail invitations to all participants.
- Similarly, the survey was configured to be rendered inactive after a two week period, on April the 8th 2008 at 6.00am DST (08-04-2008 @ 6.00 DST).
- Only one user account was used; its credentials only known to the researcher.

4.2. Responses rates

Forty-eight (48) individual participants were sent e-mail invitations to participate in this research representing a response rate of 77.08% (n=37, s=48)[1].

The response rate was calculated using the following formula:

Total response rate = [(37/48)*100]

Table 4-1 Responses rates

Number of participants	48
Number of respondents	37
Number of non-respondents	11
Total responses rates	77.08%

[1] 'n' represents the number of respondents; 's' represents the total sample si1ze.

5. Data analysis

5.1. Descriptive and inferential statistics

This section reports on the data collected through the survey construct seen in Appendix A. The data is organised per question; each question carries its own descriptive statistical analysis performed in Microsoft Excel 2003 and the Statistical Software Package for the Social Sciences (SPSS) version 15.0 for Microsoft Windows ("SPSS", 2008) as well as the validation of the hypothesis through inferential statistics (using a *one-sample t-test*)

Section 3.6 defined the two types of ordinal scales used by the participants for question 1 the one used for question 2 to 11. In order to use that data with SPSS a conversion from an ordinal scale into a numerical value understood by the software was necessary. The ordinal scale used in question 1 remains unchanged as it had already ranked the data by order of priority through a numerical assignment:

Table 5-1 Self-rated knowledge scale

No knowledge	0
Low knowledge	1
Some knowledge	2
Average knowledge	3
Good knowledge	4
Expert knowledge	5

The rating scale used for questions 2 to 11 was designed as a percentage range (Section 3.6). For analysis purposes this numeric scale was converted into an interval as follows:

Table 5-2 Risk rating scale

Very low	0-20%	1

Low	21-40%	2
Medium	41-60%	3
High	61-80%	4
Very high	81-100%	5

Those values are the ones visible on the X-axis of each histogram figure.

Question 1

Figure 5-1 depicts (in frequency histogram format) the self-rated information security knowledge score for each participant in the sample. The mean score, a measure of the "middleness" of the entire set of data for question 1, was 3.49 (μ1-1=3.49) and Standard Deviation, which indicates the degree to which scores are clusters or spread out in a distribution, was 0.989 (SD=0.989) describing low dispersion. The most commonly occurring score (the mode) was four (4).

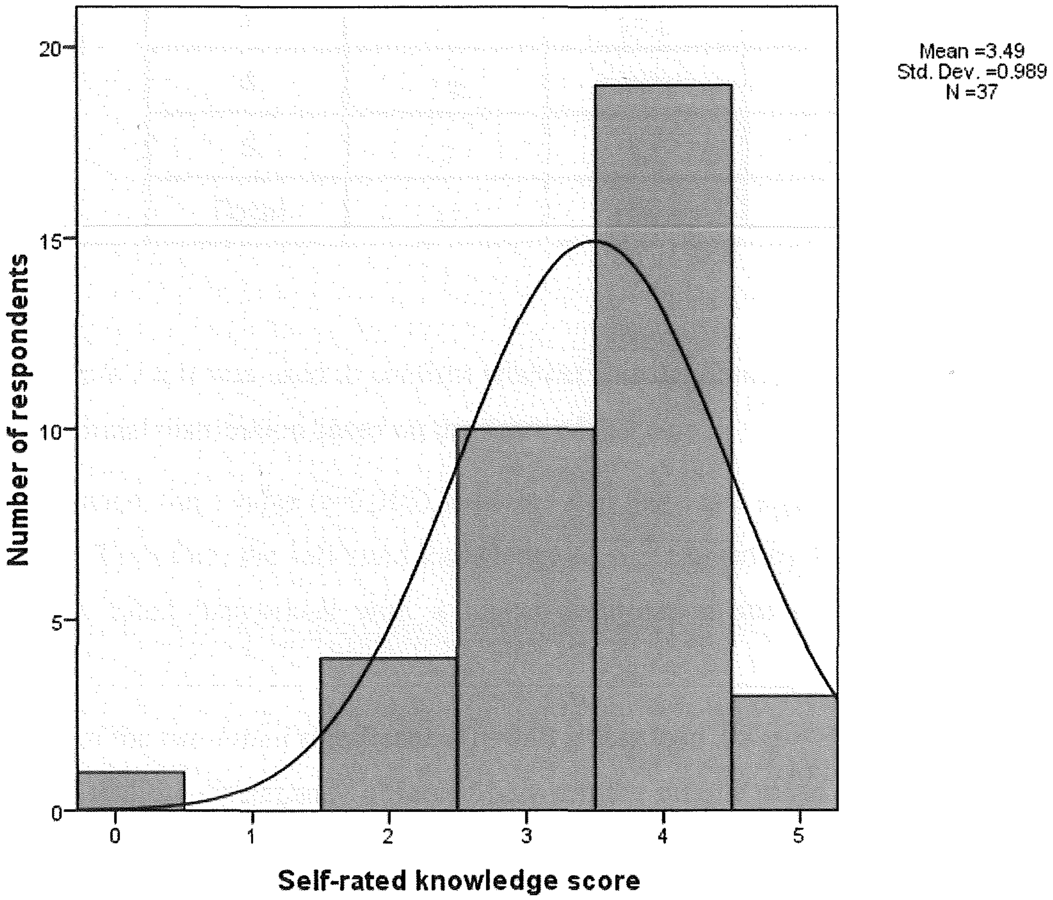

Figure 5-1 Histogram of the Self-rated knowledge score

Cumulative frequency table data for the histogram revealed that fifty nine point five percent (59.5%) of respondents rated their knowledge as 'Good' (4) or 'Very Good' (5). Only thirteen point five percent (13.5%) of respondents rated their knowledge as 'Poor' (2) or 'Very Poor' (1). Inspection of the histogram shows that relative to the normal distribution, the distribution is negatively skewed (-1.325), whilst the kurtosis (3.065), a reference to how flat or peaked a distribution is, demonstrates stronger performance scores towards the higher end of the scale with minimal dispersion amongst the scores in this particular instance.

Table 5-3 Question 1 Cumulative Percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	1	2.7	2.7	2.7
	2	4	10.8	10.8	13.5
	3	10	27.0	27.0	40.5
	4	19	51.4	51.4	91.9
	5	3	8.1	8.1	100.0
	Total	37	100.0	100.0	

A *one-sample t-test* was used to confirm whether this distribution was significantly different from the normal distribution based on the score of 2.5[2]

In this instance, the *t* value (t=6.065) indicates that there is a significant difference between the means. Therefore, the self-rated knowledge as represented by Figure 5-1 shows that the participants rated themselves with a higher information security knowledge level than average.

The value of the *two-tailed* significance (p=0.0) is less than .05 ($p < .05$). This indicates that it is unlikely that the discrepancy observed between the sample mean and average distribution mean (2.5) is due to a coincidence arising from random sampling.

This difference is only representative of the trend found from the self-rated knowledge scores. Those scores are not being compared to any other value unlike the ones for questions 2 to 11, which were compared to the values present in the information security surveys.

[2] '2.5' represents the average or 'middle' score for Question one's distribution given the chosen scale (0-5).

Question 2

Figure 5-2 depicts each respondent's opinion of the likelihood of virus attacks. The mean score (μ1-2=3.46) of the entire set of data for question 2 and Standard Deviation (SD=0.9) describe a low dispersion with a strong clustering of respondents around the mean.

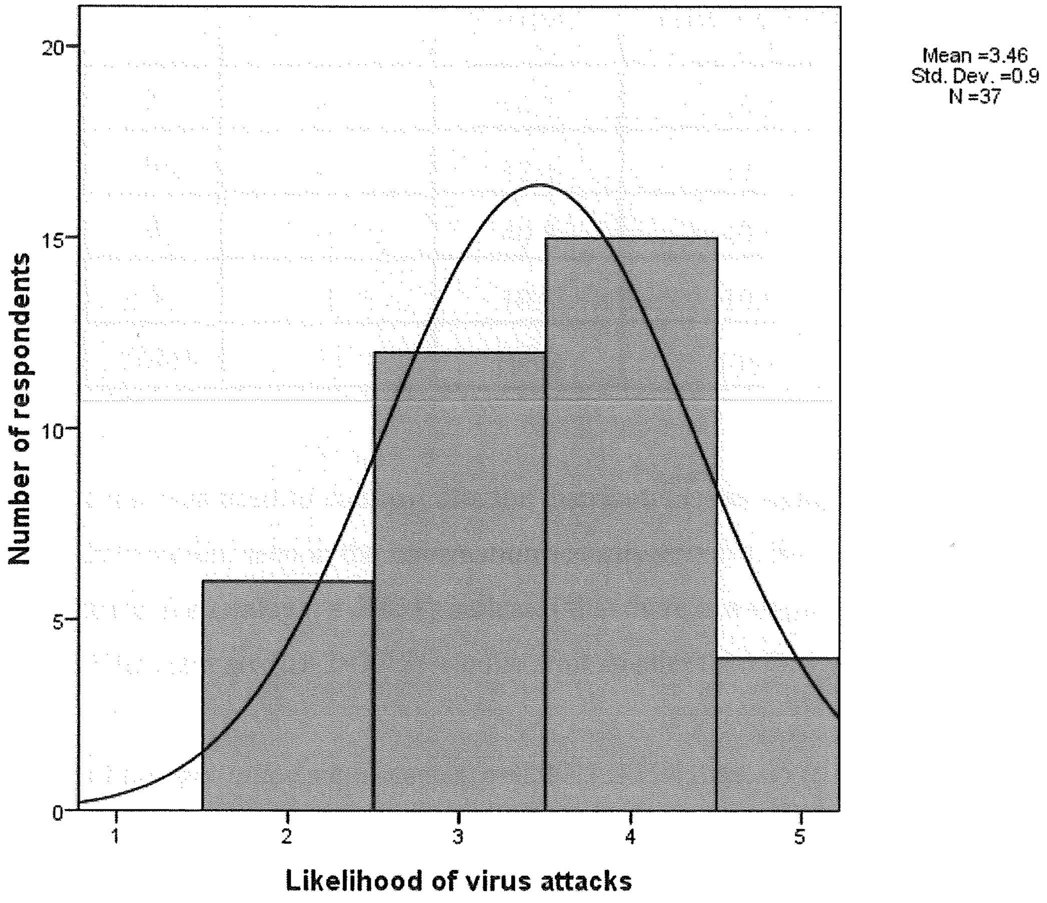

Figure 5-2 Histogram of the likelihood of virus attacks

Inspection of the frequency distribution shows that fifty one point three percent (51.3%) of respondents rated the likelihood of virus attacks as 'High' (4) or 'Very High' (5). Forty-eight point six percent (48.6%) of the respondents rated the likelihood of this risk as 'Medium' (3) or 'Low' (2). No answers were recorded for the 'Very Low' (1) rating.

The histogram and frequency table data suggest that seventy two point nine percent (72.9%) of respondents perceive the likelihood of virus attacks to be 'Medium' (3) or 'High' (4).

Inspection of the histogram shows that relative to the normal distribution, the distribution is negatively skewed (-0.114), whilst the kurtosis (-0.693) demonstrates some marginally stronger performance scores towards the higher end of the scale with minimal dispersion amongst the scores in this particular instance.

Table 5-4 Question 2 Cumulative Percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	6	16.2	16.2	16.2
	3	12	32.4	32.4	48.6
	4	15	40.5	40.5	89.2
	5	4	10.8	10.8	100.0
	Total	37	100.0	100.0	

A sample *t-test* was used to confirm that the distribution was significantly different from the reference distribution, seen in the information security surveys, based on the 'High' score (4). In this instance, the t value (t=-3.651) indicates that there is a significant difference between the means. The estimate (μ1-2=3.46) is quite a bit smaller than the hypothesized value (μ0-2=4).

The value of the *two-tailed* significance (p=0.001) is less than .05 ($p < .05$), which indicates significance. Also it so indicates that it is unlikely that the discrepancy observed between the sample mean and hypothesised mean is due to a coincidence arising from random sampling. Therefore the *one-sample t-test* output enables the rejection of the hypothesis that the mean score from the Information Security surveys for question 2 is equal to the one obtained through the online survey. A gap does exist.

Question 3

Figure 5-3 depicts the respondents' opinions of the likelihood of self-propagating malware infection (virus or worm). The data display a negatively skewed (-0.181) and kurtosis (-0.911) characteristic of some stronger performance scores towards the higher end of the scale with minimal dispersion amongst the scores in this particular instance.

The mean score (μ1-3=3.38) of the entire set of data for question 3 and Standard Deviation (SD=0.924) also describe a low dispersion with a strong clustering around the mean.

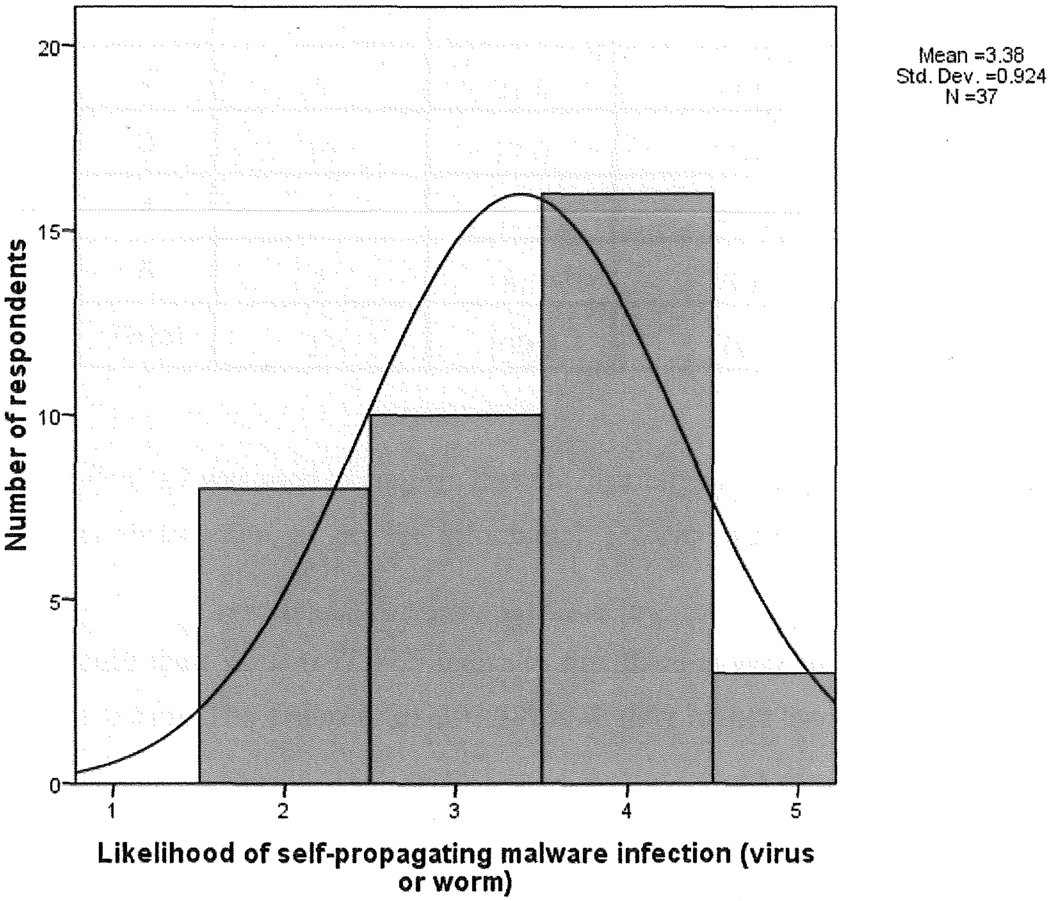

Figure 5-3 Histogram of the likelihood of self-propagating malware infection

Inspection of the frequency distribution shows that seventy point two percent (70.2%) of respondents rated the likelihood of malware infection as 'Medium' (3) or 'High' (4).

Fifty one point three percent (51.3%) of respondents rated the likelihood of malware infection as 'High' (4) or 'Very High' (5). Forty-eight point six (48.6%) of the respondents rated the likelihood of this risk as 'Medium' (3) or 'Low' (2). No answers were recorded for the 'Very Low' (1) rating.

The histogram and frequency table data suggest that seventy point two percent (70.2%) of respondents perceive the likelihood of virus attacks being 'Medium' (3) or 'High' (4).

Table 5-5 Question 3 Cumulative Percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	8	21.6	21.6	21.6
	3	10	27.0	27.0	48.6
	4	16	43.2	43.2	91.9
	5	3	8.1	8.1	100.0
	Total	37	100.0	100.0	

A *one-sample t-test* was used to confirm that the distribution was significantly different from the reference distribution, as per the information security surveys based on the 'Medium' score (3).

In this instance the *t* value (t=2.492) indicates that there is very little significant difference between the means. The estimate (μ1-3=3.38) is slightly higher than the hypothesized value (μ0-3=3).

The value of the *two-tailed* significance (p=0.378) is more than .05 ($p > .05$). This data does not provide a basis to conclude that the overall mean differs from the hypothetical value (μ0-3=3). This is not the same as saying that the true mean equals the hypothetical mean value either. Those data only indicate low significance.

Therefore, the *one-sample t-test* output enables the rejection of the hypothesis that the mean score from the information security surveys for question 3 significantly differs from the one obtained through the online survey. The existence of a gap is minimal and, in accordance with our rating scale, is not significant enough to highlight a gap.

Question 4

Figure 5-4 and summary statistics describe a slightly negative skewed distribution (-0.199) with a strong clustering of responses around the mean (μ1-4=3.22). Those measures are typical of strong performance scores. towards the higher end of the scale. The standard deviation score (SD=0.886) indicates minimal dispersion of data around the mean.

Additionally, the median (3) and standard deviation (SD=0.886) describe a strong central tendency around the mean.

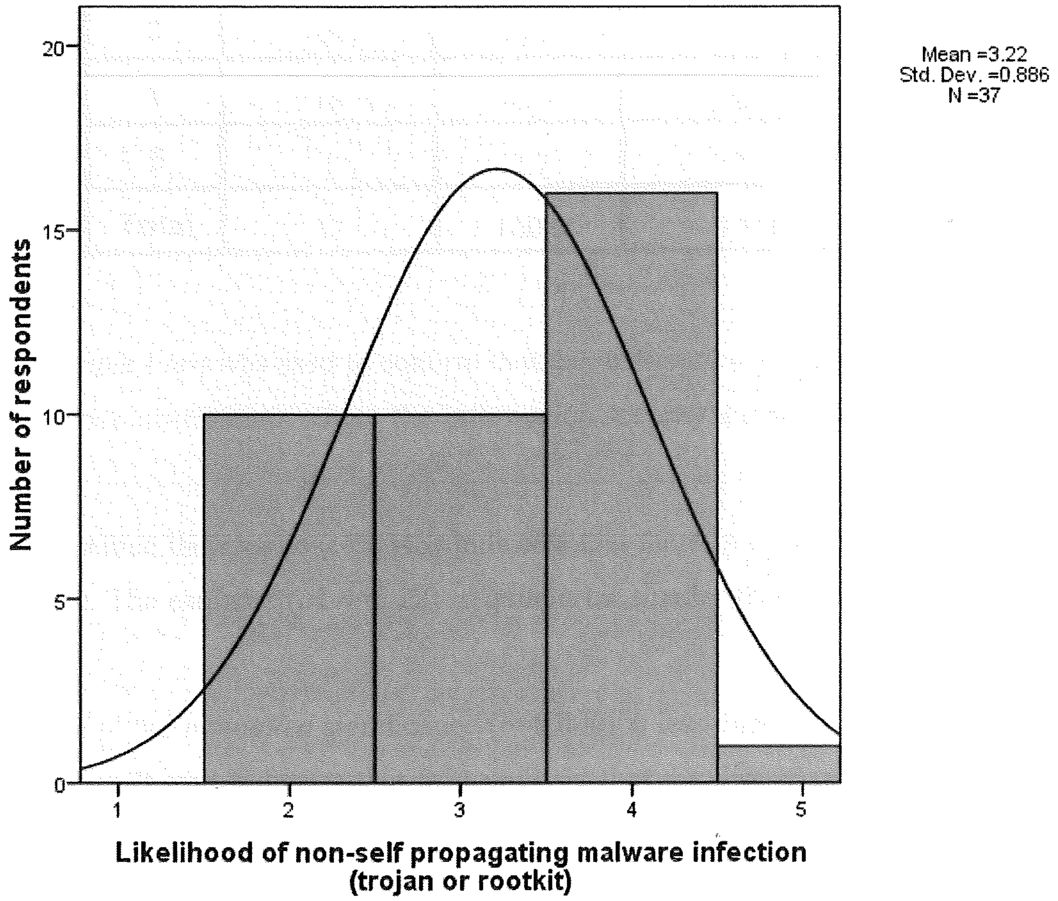

Figure 5-4 Histogram of the likelihood of non-self propagating malware infection

Investigation of cumulative frequencies shows that fifty four point nine percent (54.9%) of respondents have rated the likelihood of this risk instance as 'low' (2) or 'medium' (3), whilst forty five point nine percent (45.9%) of them rated it 'high' (4) or 'very high' (5). No responses for 'very low' (1) were recorded.

Despite the most frequently occurring result, the mode (4), being 'High' with forty three point two percent (43.2%), the median (3) and the mean (3.22) provide different measure of the central tendency for this distribution.

Table 5-6 Question 3 Cumulative Percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	10	27.0	27.0	27.0
	3	10	27.0	27.0	54.1
	4	16	43.2	43.2	97.3
	5	1	2.7	2.7	100.0
	Total	37	100.0	100.0	

A *one-sample t-test* was used to confirm that the distribution was significantly different from the reference distribution, as per the information security surveys based on the 'Low' score (2).

In this instance the t value (t=8.348) indicates that there is a significant difference between the means. The estimate (μ1-4=3.22) is quite a bit smaller than the hypothesized value (μ0-4=2).

The value of the *two-tailed* significance (p=0.000) is less than .05 ($p < .05$), which indicates significance. It also indicates that it is unlikely that the discrepancy observed between the sample mean and hypothesised mean is due to a coincidence arising from random sampling.

Therefore the *one-sample t-test* output enables the rejection of the hypothesis that the mean score from the information security surveys for question 4 is equal to the one obtained through the online survey. A gap does exist.

Question 5

Figure 5-5 depicts each the respondent's opinion of the likelihood of Denial of Service attacks. The mode (2) provides information on the most frequently occurring score, in this case 'Low' (2). The frequency distribution reveals that forty five point nine percent (45.9%) of the respondents rated the likelihood of this risk as 'Low' (2) or 'Very Low' (1). Only twenty four point three percent (24.3%) rated it as "High' (4) or 'Very High" (5).

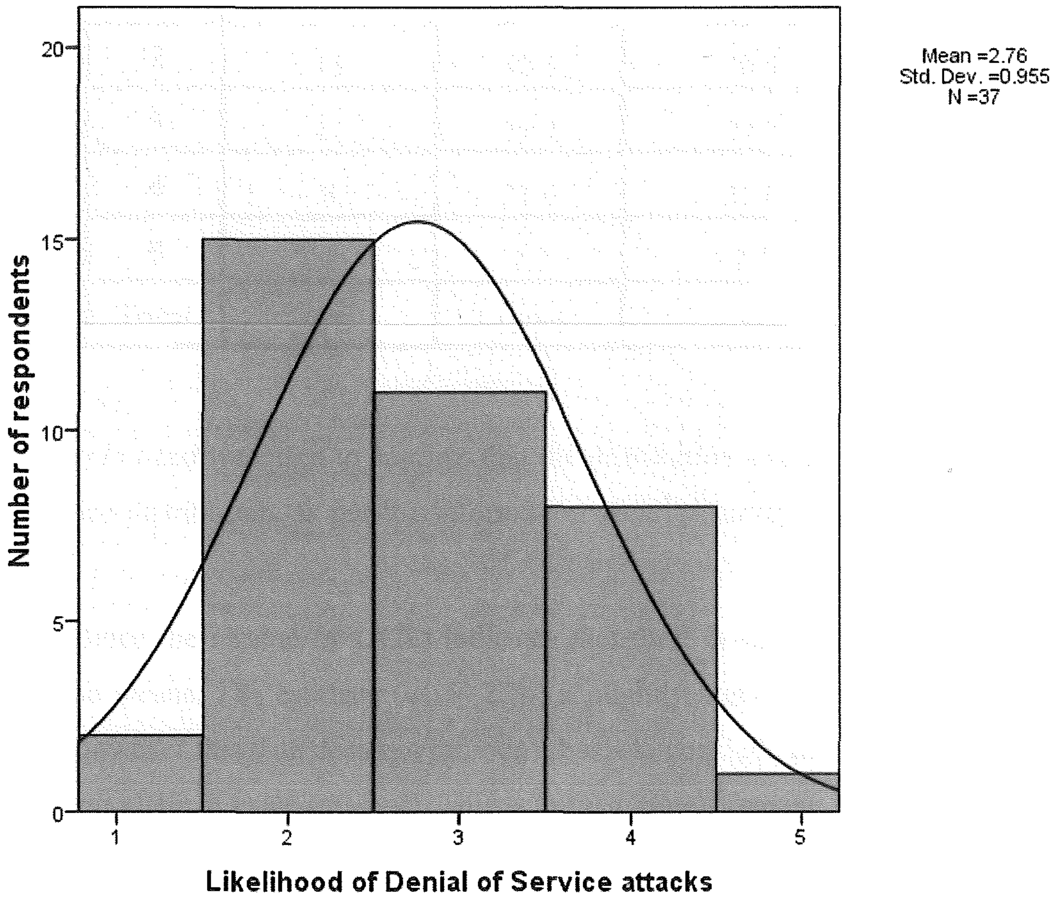

Figure 5-5 Histogram of the likelihood of denial of service attacks

The mean (μ1-5=2.76) and median (3) provide an indication of the central tendency distribution. The median resides eight point seven percent away from the mean (8.7%). The distribution is positively skewed (0.320) with a slightly negative kurtosis (-0.563) that demonstrates stronger scores towards the front-end of the scale. The standard deviation

(SD=0.955) indicates little variation around the mean's distribution. Inspection of the frequency table reveals that seventy point two percent (70.2%) of the respondents considered the likelihood of this risk as 'Low' (2) or 'Medium' (3).

Table 5-7 Question 5 Cumulative Percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	2	5.4	5.4	5.4
	2	15	40.5	40.5	45.9
	3	11	29.7	29.7	75.7
	4	8	21.6	21.6	97.3
	5	1	2.7	2.7	100.0
	Total	37	100.0	100.0	

A *one-sample t-test* was used to confirm that the distribution was significantly different from the reference distribution, as per the information security surveys based on the 'Low' score (2).

In this instance the *t* value (*t*=4.822) indicates that there is very little significant difference between the means. The estimate (μ1-5=2.76) is slightly higher than the hypothesized value (μ0-5=2), however this data does not provide a basis to conclude that the overall mean differs from the hypothetical value (μ0-5=2). This is not the same as saying that the true mean equals the hypothetical mean value either. Those data only indicate low significance.

The value of the *two-tailed* significance (p=0.000) is less than .05 ($p < .05$). This indicates that it is unlikely that the discrepancy observed between the sample mean and hypothesised mean is due to a coincidence arising from random sampling.

Therefore the *one-sample t-test* output enables the rejection of the hypothesis that the mean score from the information security surveys for question 5 significantly differs from the one obtained through the online survey. The existence of a gap is minimal and, in accordance to our rating scale, is not significant enough to highlight a gap.

Question 6

Figure 5-6 depicts each respondent's opinion of the likelihood of theft or breach of proprietary or confidential information.

The distribution displays a central tendency. The data display a negatively skewed (-0.767) with positive kurtosis (0.388). The mean score (μ1-6=3.3) and median (4) of the entire set of data for question 6 and Standard Deviation (SD=1.102) describe a low dispersion with the majority of the observations being contained within one standard deviation of the mean.

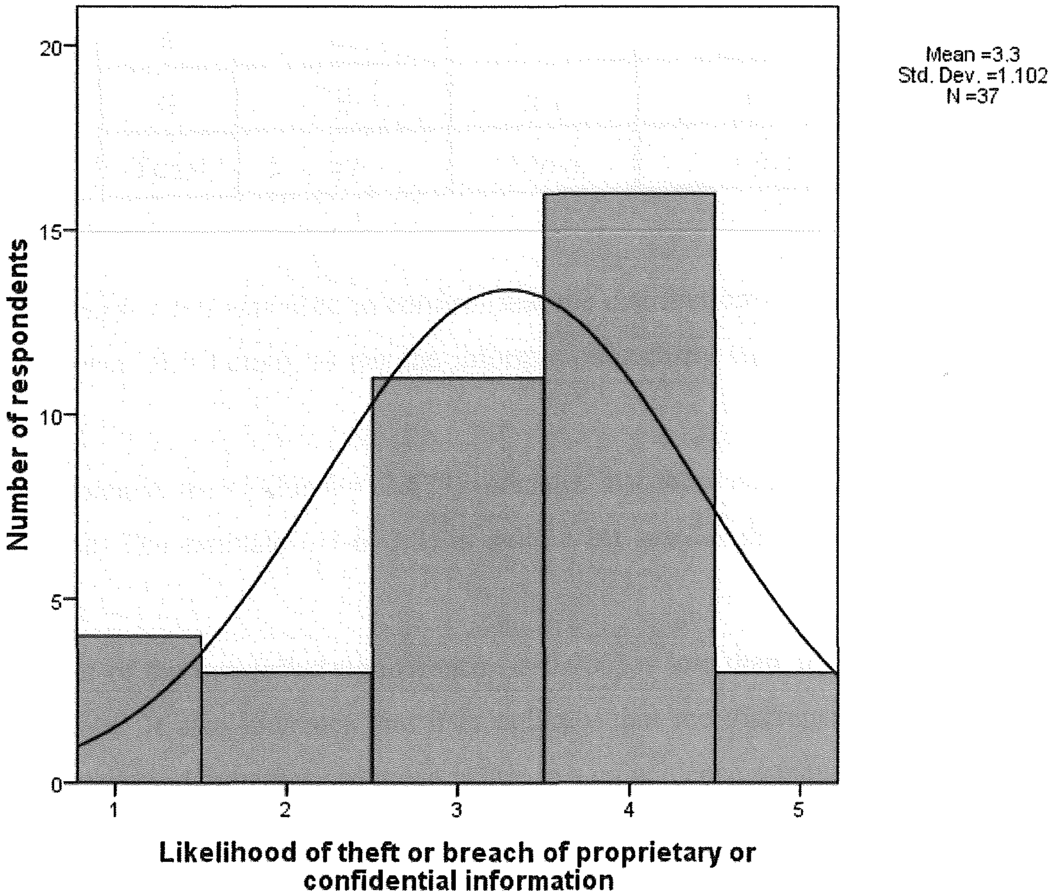

Figure 5-6 Histogram of the likelihood of theft or breach of proprietary or confidential information

Inspection of the frequency distribution shows that eighteen point nine percent (18.9%) of respondents rated the likelihood of theft or breach of proprietary or confidential information as 'Very Low' (1) or 'Low' (2) whilst seventy two point nine percent (72.9%) of the respondents rated the likelihood of this risk as 'Medium' (3) or 'High' (4).

Table 5-8 Question 6 Cumulative Percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	4	10.8	10.8	10.8
	2	3	8.1	8.1	18.9
	3	11	29.7	29.7	48.6
	4	16	43.2	43.2	91.9
	5	3	8.1	8.1	100.0
	Total	37	100.0	100.0	

A *one-sample t-test* was used to confirm that the distribution was significantly different from the reference distribution, as per the information security surveys based on the 'Very Low' score (1).

In this instance, the *t* value (*t*=12.679) indicates that there is a significant difference between the means. The estimate (μ1-6=3.3) is quite a bit smaller than the hypothesized value (μ0-6=1).

The value of the *two-tailed* significance (*p*=0.000) is less than .05 ($p < .05$), which indicates significance. It also indicates that it is unlikely that the discrepancy observed between the sample mean and hypothesised mean is due to a coincidence arising from random sampling.

Therefore the *one-sample t-test* output enables the rejection of the hypothesis that the mean score from the information security surveys for question 6 is equal to the one obtained through the online survey. Based on the t value a significant gap exists.

Question 7

Figure 5-7 depicts the likelihood of System penetration by an outsider as seen by each participant in the sample. The mean (μ1-7=2.7) of the entire set of data for question 7 and Standard Deviation (SD=0.777) indicates a minimal dispersion.

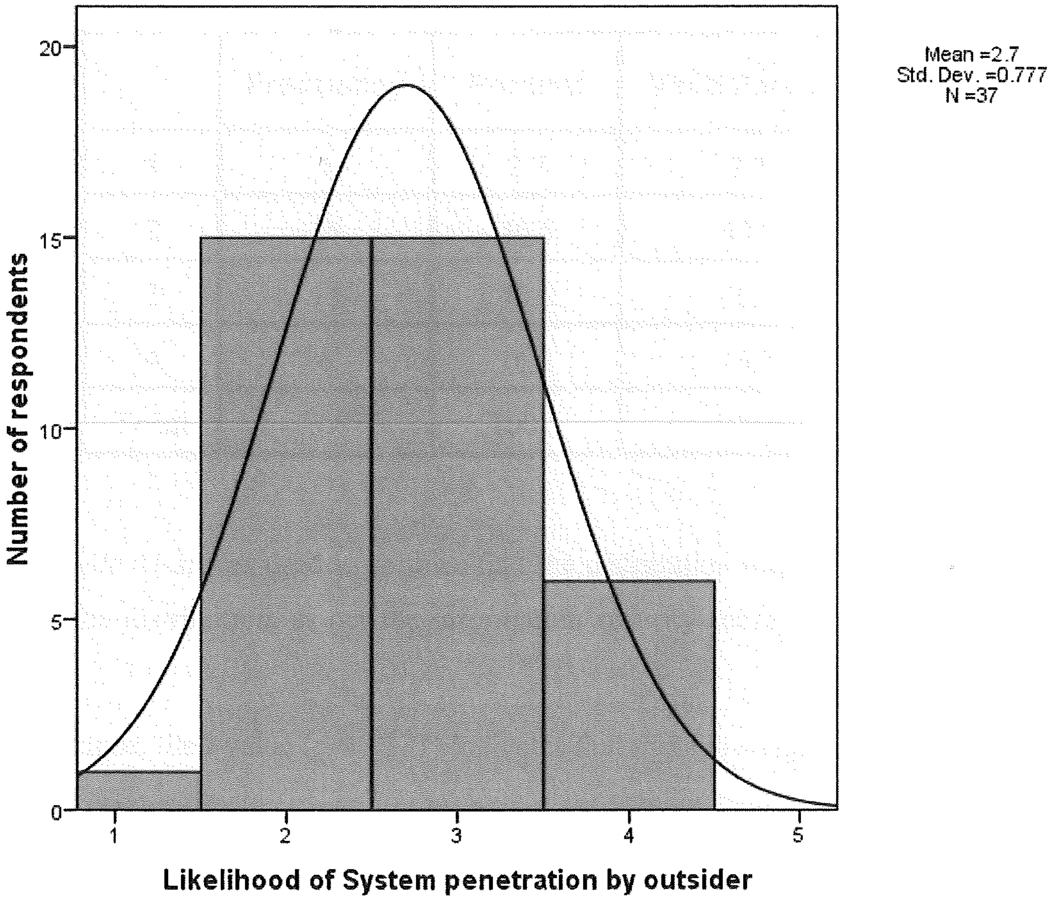

Figure 5-7 Histogram of the likelihood of system penetration by outsider

Inspection of the histogram and frequency table reveals a strong clustering of responses around the mean. The mode indicates that two scores (2 and 3) have an equal distribution, therefore they have the same frequency. This is re-enforced by the median (3) and Standard Deviation (SD=0.777) that describe a strong central tendency around the mean.

Eighty one percent (81%) of the respondents perceiving the likelihood associated with the risk 'System penetration by outsiders' as 'Low' (2) or 'Medium' (3). Fifty six point seven percent

(56.7%) of the respondents rated this risk as 'Medium' (3) or 'High' (4) with no responses were recorded for 'Very High' (5).

Inspection of the histogram shows a positively skewed (-0.114) distribution whilst the kurtosis (-0.639) demonstrates strong performance scores towards the higher end of the scale.

Table 5-9 Question 7 Cumulative Percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1	2.7	2.7	2.7
	2	15	40.5	40.5	43.2
	3	15	40.5	40.5	83.8
	4	6	16.2	16.2	100.0
	Total	37	100.0	100.0	

A *one-sample t-test* was used to confirm that the distribution was significantly different from the reference distribution, as per the information security surveys based on the 'Very Low' score (1).

In this instance, the *t* value (t=13.331) indicates that there is a significant difference between the means. The estimate (μ1-7=2.7) is quite a bit smaller than the hypothesized value (μ0-7=1).

The value of the *two-tailed* significance (p=0.000) is less than .05 ($p < .05$), which indicates significance. It also indicates that it is unlikely that the discrepancy observed between the sample mean and hypothesised mean is due to a coincidence arising from random sampling.

Therefore the *one-sample t-test* output enables the rejection of the hypothesis that the mean score from the information security surveys for question 7 is equal to the one obtained through the online survey. Based on the t value, a significant gap exists.

Question 8

Figure 5-8 depicts the respondent's opinion of the likelihood of Laptop theft. The data display a negatively skewed (-0.504) and kurtosis (-0.661) characteristic of stronger performance scores towards the higher end of the scale.

The mean score (μ1-8=3.73) of the entire set of data for question 8 and Standard Deviation (SD=0.99) also describe a low dispersion.

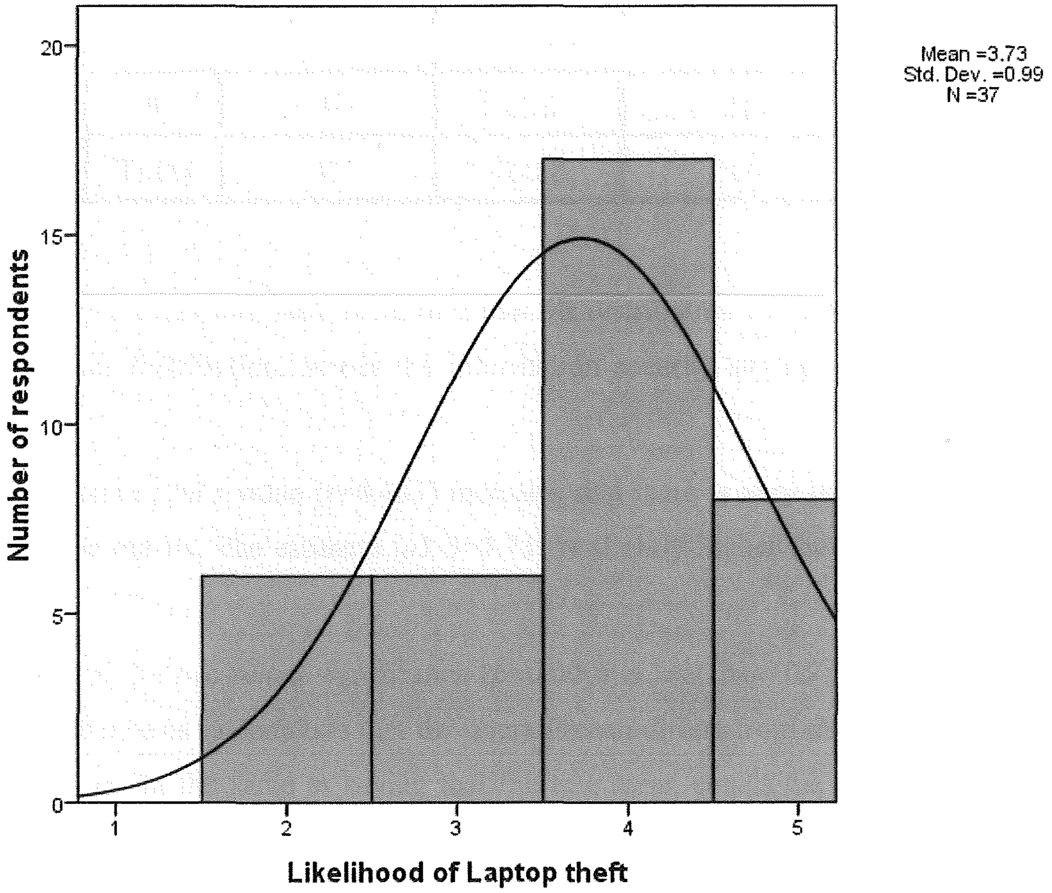

Figure 5-8 Histogram of the likelihood of laptop theft

Inspection of the frequency distribution shows that the mode (4) occurs with a frequency (17) that is more than double the ones obtained by other scores.

Sixty seven point five percent (67.5%) of the respondents rated the likelihood of this risk as 'High' (4) or 'Very High' (5).

No scores were recorded for 'Very Low' (1). In addition, thirty two point four percent (32.4%) of the respondents perceived the likelihood of this risk as 'Low' (2) or 'Medium' (3).

Table 5-10 Question 8 Cumulative Percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	6	16.2	16.2	16.2
	3	6	16.2	16.2	32.4
	4	17	45.9	45.9	78.4
	5	8	21.6	21.6	100.0
	Total	37	100.0	100.0	

A *one-sample t-test* was used to confirm that the distribution was significantly different from the reference distribution, as per the information security surveys based on the 'Medium' score (3).

In this instance, the *t* value (*t*=4.483) indicates that there is very little significant difference between the means. The estimate (μ1-3=3.73) is slightly higher than the hypothesized value (μ0-3=3).

The value of the *two-tailed* significance (*p*=0.000) is less than .05 ($p < .05$). This data does not provide a basis to conclude that the overall mean differs from the hypothetical value (μ0-3=3). This is not the same as saying that the true mean equals the hypothetical mean value either. Those data only indicate low significance.

The *one-sample t-test* output therefore enables the rejection of the hypothesis that the mean score from the information security surveys for question 8 significantly differs from the one obtained through the online survey. The existence of a gap is minimal and in accordance to our rating scale, is not significant enough to highlight a gap.

Question 9

Figure 5-9 depicts the respondent's opinion of the likelihood of Insider abuse of Internet access, e-mail or computer resources.

The data set display a negatively skew (-0.761) and kurtosis (-0.669) that depicts a relatively asymmetric distribution with heavy distribution towards the highest scores. Similarly, the median (5) and Standard Deviation (SD=0.895) are consistent with the previous observation.

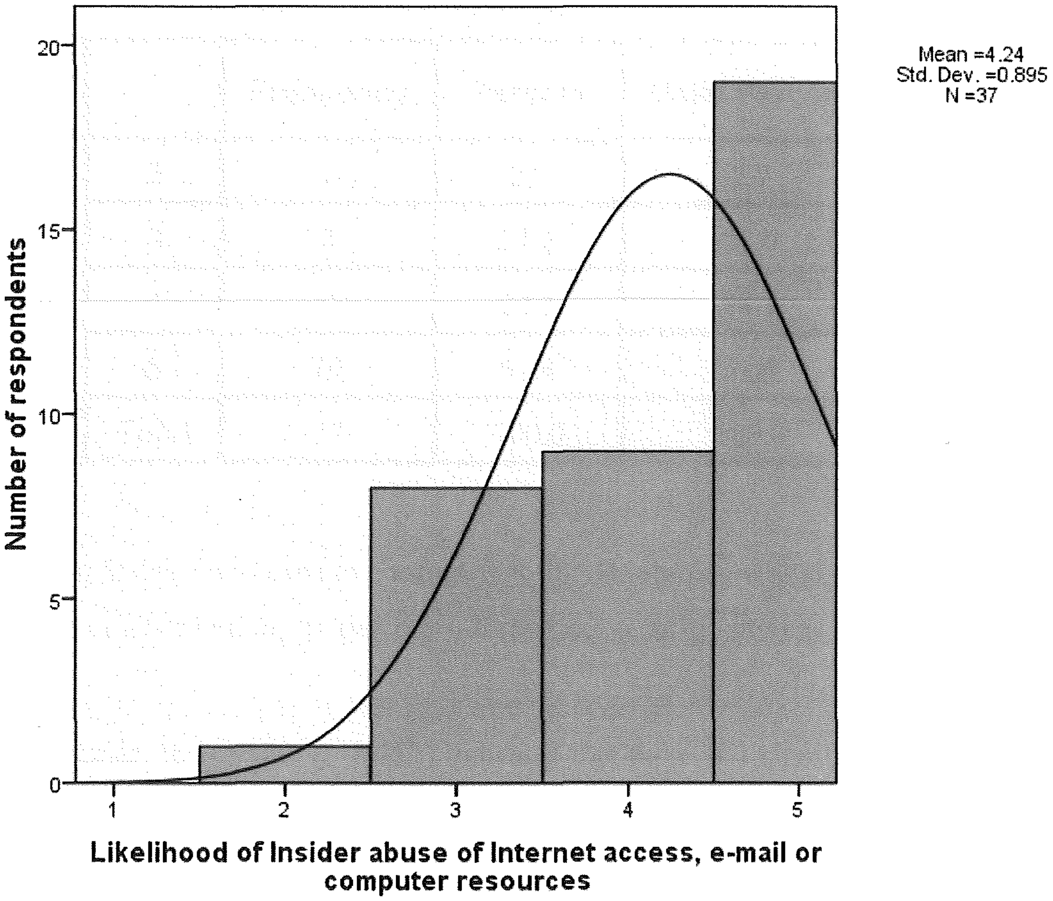

Figure 5-9 Histogram of the likelihood of insider abuse of Internet access, e-mail or computer resources

Inspection of the frequency distribution shows that the mode (5), associated with fifty one point four percent (51.4%) of all responses, occurs with a frequency (19) that is more than double the ones obtained by other scores. In other words, 51.4% of all respondents perceive the likelihood of the risk 'Insider abuse of Internet access, e-mail or computer resources' as 'Very High'.

A cumulative seventy five point eight percent (75.8%) perceived the likelihood of the same risk as 'High' (4) or 'Very High' (5). Comparatively, twenty four point three percent (24.3%) of the respondents recorded the likelihood of this risk as 'Low' (2) or 'Medium' (3).

No answers were recorded for 'Very Low' (1).

Table 5-11 Question 9 Cumulative Percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	1	2.7	2.7	2.7
	3	8	21.6	21.6	24.3
	4	9	24.3	24.3	48.6
	5	19	51.4	51.4	100.0
	Total	37	100.0	100.0	

A *one-sample t-test* was used to confirm that the distribution was significantly different from the reference distribution, as per the information security surveys based on the 'Medium' score (3).

In this instance, the t value (t=8.453) indicates that there is a significant difference between the means. The estimate (μ1-9=4.24) is quite a bit smaller than the hypothesized value (μ0-9=3).

The value of the *two-tailed* significance (p=0.000) is less than .05 (p < .05), which indicates significance. It also indicates that it is unlikely that the discrepancy observed between the sample mean and hypothesised mean is due to a coincidence arising from random sampling.

Therefore the *one-sample t-test* output enables the rejection of the hypothesis that the mean score from the information security surveys for question 9 is equal to the one obtained through the online survey. A gap does exist.

Question 10

Figure 5-10 depicts likelihood of the risk 'Unauthorized access to information by insider' as rated by each participant in the sample.

The mean score (μ1-10=3.49) for the entire set of data for question 10 and Standard Deviation (SD=0.901) describes a low dispersion.

The distribution displays a strong central tendency with a negative skew (-0.679) and a positive kurtosis (0.481) demonstrating a minimal dispersion with better than average scores which are more frequent towards the higher end of the rating scale. This is seen also with the mode (4) and median (4) scores.

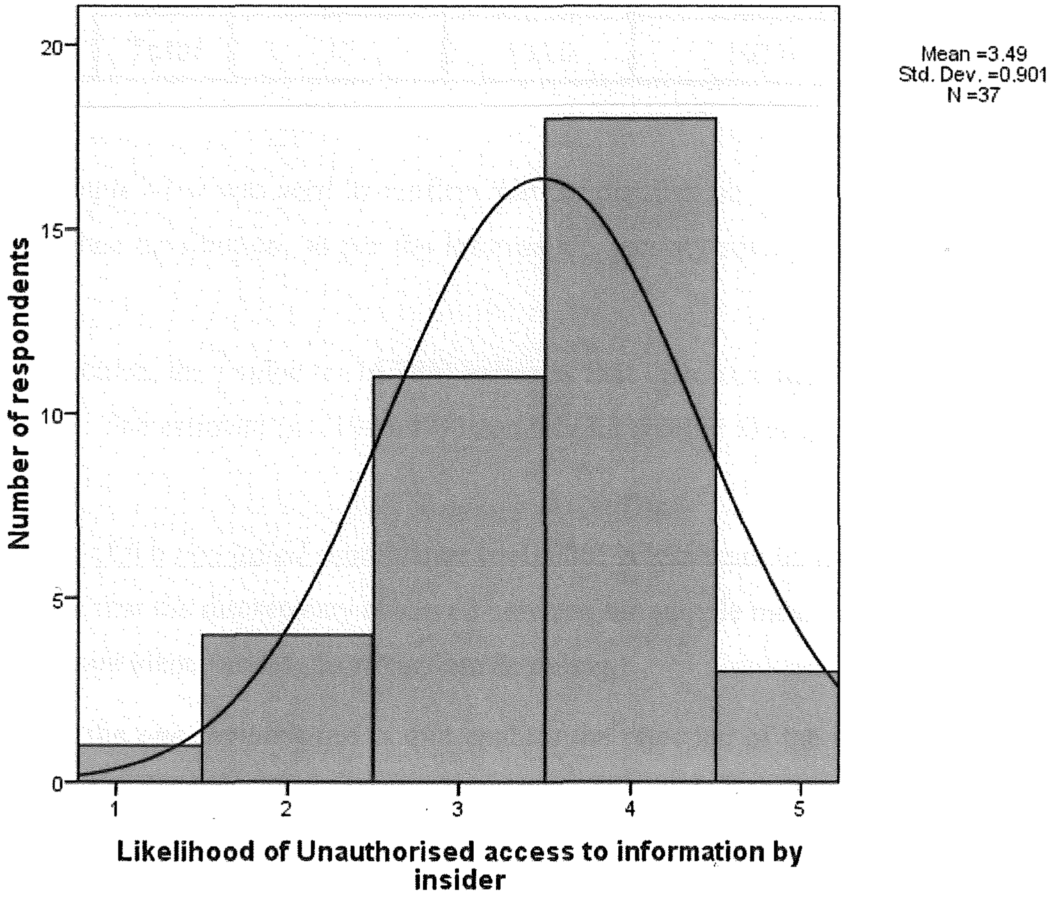

Figure 5-10 Histogram of the likelihood of unauthorised access to information by insider

Cumulative frequency table data for the histogram revealed that fifty six point seven percent (56.7%) of respondents rated the likelihood of this risk as 'High' (4) or 'Very High' (5). Forty three point two percent (43.2%) of the respondents rated the likelihood of this risk as 'Very Low' (1), 'Low' (2) or 'Medium' (3).

Table 5-12 Question 10 Cumulative Percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1	2.7	2.7	2.7
	2	4	10.8	10.8	13.5
	3	11	29.7	29.7	43.2
	4	18	48.6	48.6	91.9
	5	3	8.1	8.1	100.0
	Total	37	100.0	100.0	

A *one-sample t-test* was used to confirm that the distribution was significantly different from the reference distribution, as per the information security surveys based on the 'Very Low' score (1).

In this instance, the t value (t=16.781) indicates that there is a significant difference between the means. The estimate (μ1-10=3.49) is quite a bit smaller than the hypothesized value (μ0-10=1).

The value of the *two-tailed* significance (p=0.000) is less than .05 ($p < .05$), indicating that it is unlikely that the discrepancy observed between the sample mean and hypothesised mean is due to a coincidence arising from random sampling.

Therefore the *one-sample t-test* output enables the rejection of the hypothesis that the mean score from the information security surveys for question 10 is equal to the one obtained through the online survey. Based on the t value a significant gap exists.

Question 11

Figure 5-11 depicts each respondent's opinion of the likelihood of Website defacement.

The data display a moderately positive skew (0.265) and negative kurtosis (-0.505), indicating stronger scores towards the front-end of the scale, with a strong clustering of responses around the mean ($\mu 1$).

The mean score ($\mu 1$-11=2.46) of the entire set of data for question 11 and Standard Deviation (SD=0.869) describe a low dispersion.

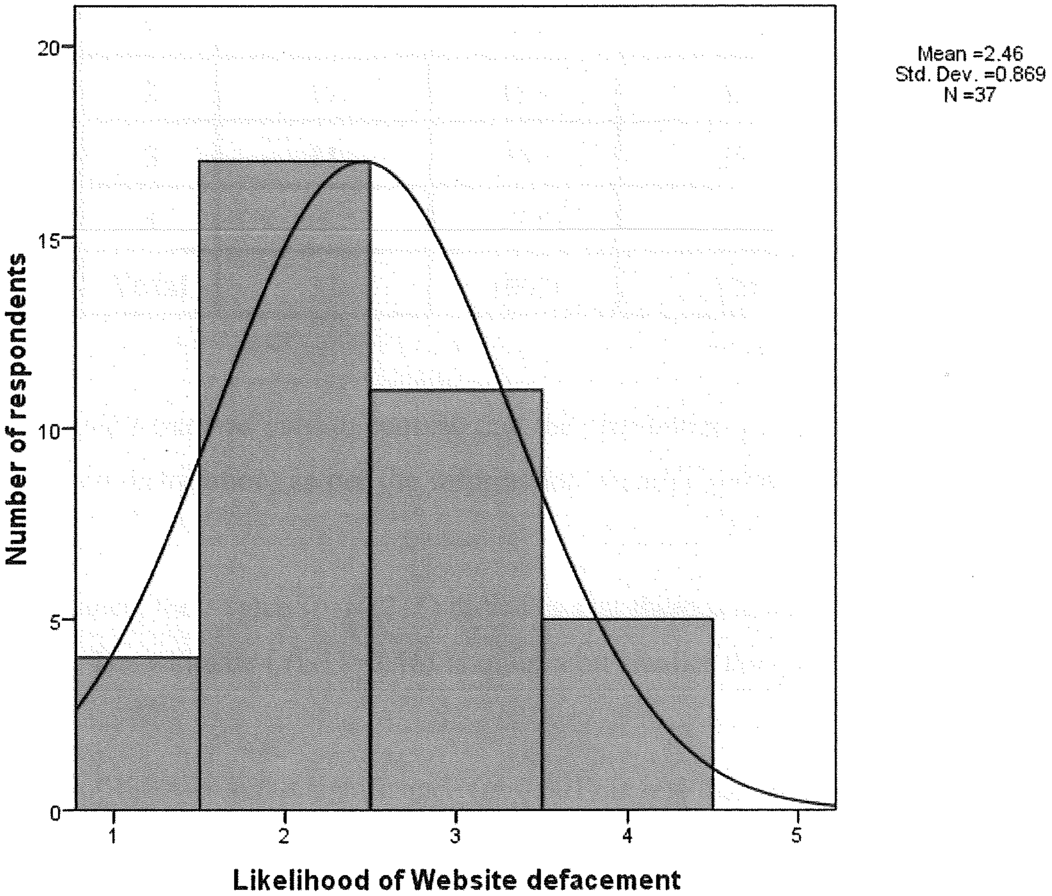

Figure 5-11 Histogram of the likelihood of website defacement

Inspection of cumulative frequencies shows that forty three point two percent (43.2%) of the respondents rated the likelihood of Website defacement as 'Medium' (3) or 'High' (4) whilst no answers were recorded for 'Very High' (5).

Seventy five point six percent (75.6%) rated the likelihood of this risk as 'Low' (2) or 'Medium' (3). Fifty six point eight percent (56.8%) of the respondents perceived the likelihood of this risk as 'Very Low' (1) or 'Low' (2).

The score occurring with the most frequency or mode is 'Low' (2) with seventeen (17) occurrences representing forty five point nine percent (45.9%) of all responses.

Table 5-13 Question 11 Cumulative Percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	4	10.8	10.8	10.8
	2	17	45.9	45.9	56.8
	3	11	29.7	29.7	86.5
	4	5	13.5	13.5	100.0
	Total	37	100.0	100.0	

A *one-sample t-test* was used to confirm that the distribution was significantly different from the reference distribution, as per the information security surveys based on the 'Very Low' score (1).

In this instance, the t value (t=10.215) indicates that there is a significant difference between the means. The estimate (μ1-11=2.46) is quite a bit smaller than the hypothesized value (μ0-10=1).

The value of the *two-tailed* significance (p=0.001) is less than .05 ($p < .05$), which indicates significance. It also indicates that it is unlikely that the discrepancy observed between the sample mean and hypothesised mean is due to a coincidence arising from random sampling.

Therefore the *one-sample t-test* output enables the rejection of the hypothesis that the mean score from the information security surveys for question 11 is equal to the one obtained through the online survey. Based on the t value a significant gap exists.

5.2. Hypothesised mean scores

Providing the results of the *one-sample t-test*, this research was able to highlight variances or gaps between the mean of the population from which the sample was drawn and the hypothesised mean as per the information security surveys. The following section uses a traffic light representation to visually highlight the gaps between the means.

Hypothesised means from the information security surveys

The hypothesised mean ($\mu 0$) originating from the information security surveys was easily calculated for each of the ten risk occurrences using the sum of individual percentage scores and divided by the total number of scores in the distribution (N=4), providing the following data.

Table 5-14 Risk occurrence percentages as per information security surveys

Selected risks (As per Appendix A)	AusCERT 2005	AusCERT 2006	CSI/FBI 2006	CSI/FBI 2007	Mean ($\mu 0$) score
Q2 - Virus attacks	64%	66%	65%	52%	**61.75%**
Q3 - Self-propagating malware infection (virus or worm)	- %	45%	- %	- %	**45%**
Q4 - Non-propagating malware infection (Trojan or rootkit)	- %	21%	- %	- %	**21%**
Q5 - Denial of Service attacks	24%	18%	25%	25%	**23%**
Q6 - Theft or breach of proprietary or confidential information	14%	14%	9%	8%	**11.25%**
Q7 - System penetration by outsider	6%	7%	15%	13%	**10.25%**
Q8 - Laptop theft	53%	58%	47%	50%	**52%**
Q9 - Insider abuse of Internet access, e-mail or computer resources	68%	62%	42%	59%	**57.75%**

Q10 - Unauthorized access to information by insider	9%	8%	32%	25%	**18.5%**
Q11 - Website defacement	8%	8%	6%	10%	**8%**

Those mean scores, shown in Table 5-15 were named in accordance with the following convention:

- The mean score for the information security surveys is μ0
- The mean score for question X as rated by the information security surveys is μ0-X

Table 5-15 Mean scores naming conventions for the information security surveys

Information security survey risks equivalent to:	Mean naming convention[3]
Question 2	μ0-2
Question 3	μ0-3
Question 4	μ0-4
Question 5	μ0-5
Question 6	μ0-6
Question 7	μ0-7
Question 8	μ0-8
Question 9	μ0-9
Question 10	μ0-10
Question 11	μ0-11

The conversion scale previously discussed in Section 3.6 were leveraged to facilitate the conversion of the mean percentage values obtained from the information security surveys into a numerical interval as follows:

[3] Question 1 for this research survey does not exist in the information security surveys therefore only Question 2-11 are being used in this section.

Table 5-16 Risk rating scale

Very low	0-20%	(light green)
Low	21-40%	(green)
Medium	41-60%	(yellow)
High	61-80%	(orange)
Very high	81-100%	(red)

Table 5-17 outlines the equivalent mean score for each risk question (Q2-11) in accordance with that scale.

Table 5-17 Hypothesised mean scores from the information security surveys

Information security survey risks equivalent to:	Very Low (0-20%)	Low (21-40%)	Medium (41-60%)	High (61-80%)	Very High (81-100%)
Question 2				μ0-2=4	
Question 3			μ0-3=3		
Question 4		μ0-4=2			
Question 5		μ0-5=2			
Question 6	μ0-6=1				
Question 7	μ0-7=1				
Question 8			μ0-8=3		
Question 9			μ0-9=3		
Question 10	μ0-10=1				

Question 11	μ0-11=1				

Hypothesised means from the online survey

Following the same process as the previous section, the mean (μ1) drawn from the online survey construct was individually named in accordance with the following convention as shown in Table 5-18

- The mean score for the online survey construct is μ1
- The mean score for question X as rated by the population in the online survey is μ1-X

Table 5-18 Mean scores naming conventions for targeted sample population

Question number	Mean naming convention[4]
Question 2	μ1-2
Question 3	μ1-3
Question 4	μ1-4
Question 5	μ1-5
Question 6	μ1-6
Question 7	μ1-7
Question 8	μ1-8
Question 9	μ1-9
Question 10	μ1-10
Question 11	μ1-11

[4] The mean value for Question 1 (μ1-1) is not relevant for the comparison of the 'gap', therefore not represented in this section.

(Q2-11).

Table 5-19 outlines the mean scores for each risk question (Q2-11).

Table 5-19 Mean scores drawn from the online survey[5]

Survey questions (As per Appendix A)	Very Low (0-20%)	Low (21-40%)	Medium (41-60%)	High (61-80%)	Very High (81-100%)
Question 2			μ1-2=3.4		
Question 3			μ1-3=3.3		
Question 4			μ1-4=3.2		
Question 5		μ1-5=2.7			
Question 6			μ1-6=3.3		
Question 7		μ1-7=2.7			
Question 8			μ1-8=3.7		
Question 9				μ1-9=4.2	
Question 10			μ1-10=3.4		
Question 11		μ1-11=2.4			

[5] Mean values (μ1-X) have been rounded to one decimal point.

5.3. *Is there a gap in people's perception of those risks?*

Research question: 'Is there a gap between people's perception of information security risks and the findings of published research of those risks?'

One of the aims of the research questions was to highlight whether or not a gap existed between people's perception of information security risks and the findings of published research of those risks. At this stage of the research, enough analysis elements have been discovered to provide an answer to this question.

Based on this research methodology, measurements, information sources and statistical analysis, the answer to this question is yes, a gap does exist between the perceptions of certain information security risks and the published research of those risks.

(Q2-11).

Table 5-19 indicates dispersion in the way the answers are displayed. Table 5-20 shows one set of data superimposed on the other to provide an easy and visual location of the overlaps.

It also becomes clear that question 3, 5 and 8 do not present a gap significant enough, in accordance to the rating scales used by this research, to prompt a change of category.

The mean scores recorded for the remaining questions (2, 4, 6, 7, 9, 10 and 11) were, on the other hand, significantly different to present a gap between the respondents' answers and the reference mean scores obtained from the information security surveys.

Table 5-20 Superimposed mean scores (table)

Risks (As per Appendix A)	Very Low (0-20%)	Low (21-40%)	Medium (41-60%)	High (61-80%)	Very High (81-100%)
Q2 - Virus attacks			μ1-2=3.4	μ0-2=4	
Q3 - Self-propagating malware infection (virus or worm)			μ0-3=3 μ1-3=3.4		
Q4 - Non-propagating malware infection (Trojan or rootkit)		μ0-4=2	μ1-4=3.2		
Q5 - Denial of Service attacks		μ0-5=2 μ1-5=2.7			
Q6 - Theft or breach of proprietary or confidential information	μ0-6=1		μ1-6=3.3		
Q7 - System penetraion by outsider	μ0-7=1	μ1-7=2.7			
Q8 - Laptop theft			μ0-8=3 μ1-8=3.7		
Q9 - Insider abuse of Internet access, e-mail or computer resources			μ0-9=3	μ1-9=4.2	
Q10 - Unauthorized access to information by insider	μ0-10=1		μ1-10=3.4		
Q11 - Website defacement	μ0-11=1	μ1-11=2.4			

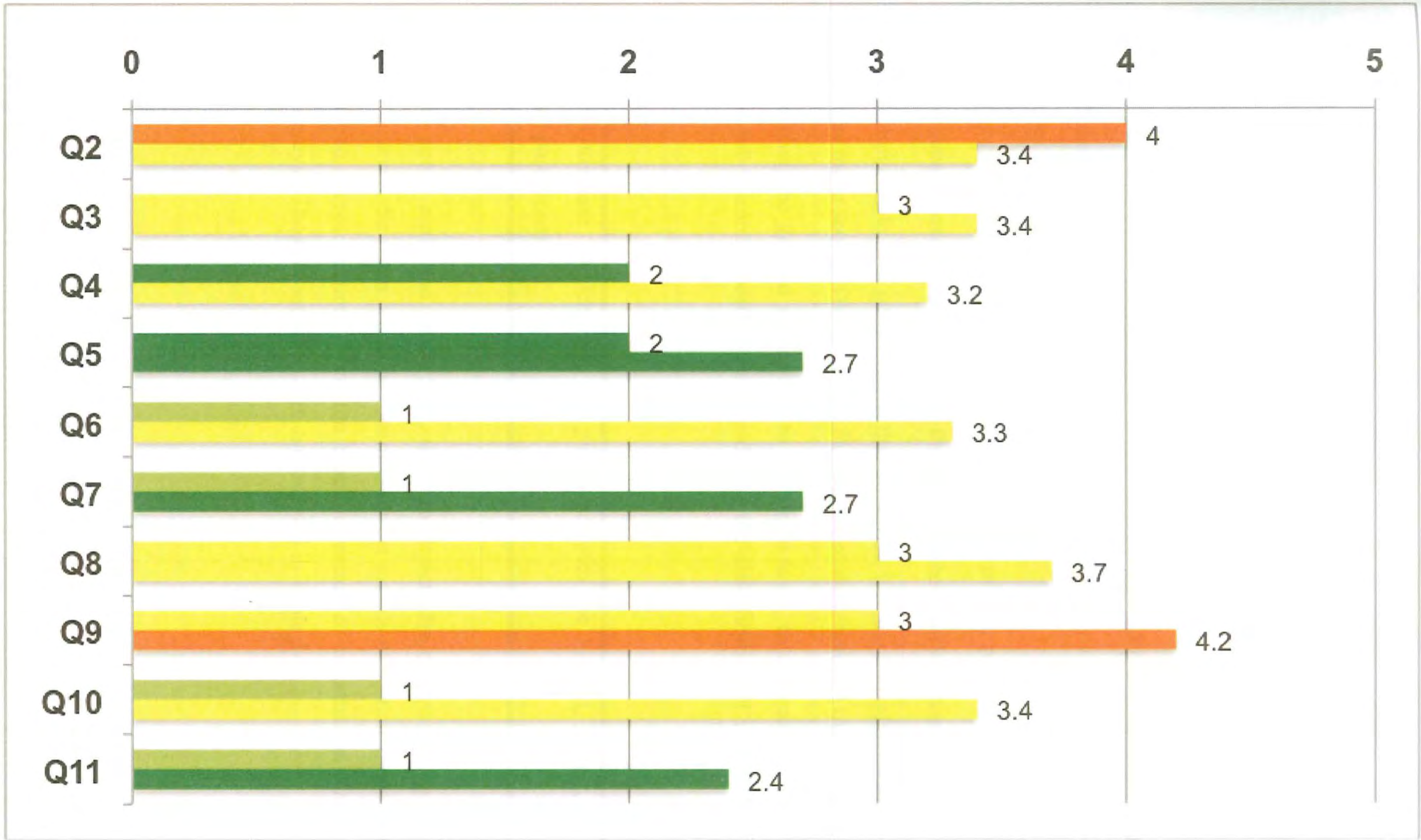

Figure 5-12 Superimposed mean scores (graph): depicts the risk perception gaps using the same colour scheme as Table 5-17

5.4. *Measure of the gap*

In accordance with the rating scales (Section 3.6), the samples were categorised into groups of different gap strength. Each group were associated with an ordinal metric value defined by its *variance* with the reference answer from the surveys '$\mu 0$' (as per Table 5-17).

As a reminder the proposed scale was defined as follows:

- High gaps – participant answers have a variance of +/- 3 points on the five-point-scale.
- Medium gaps – participant answers have a variance of +/- 2 points on the five-point-scale.
- Low gaps – participant answers have a variance of +/- 1 point on the five-point-scale.
- No gaps – participant answers are the same and have no variance

Using the above scale, it is quiet evident that questions 3, 5 and 8 display no differences with the reference mean scores. Therefore, the focus turns to the remaining questions.

By categorising the gaps in accordance to the above scale, this research effectively answer one of the sub-research question concerned with the classification of the gaps in perception variances.

Research Question: 'Which risks are perceived with the highest, lowest, or same gap amongst all participants?'

As per the above scale and Table 5-20, the data depicts the existence of gaps of various strengths between the mean scores as recorded by the sample of participants to this research and the mean scores obtained from the baseline source of information (the information security surveys). Those observations are summarised in Table 5-21.

Table 5-21 Gaps strengths and directions

Question number	Recorded gap strength	Direction of the gap
Q2 - Virus attacks	Low	$\mu0\text{-}2 > \mu1\text{-}2$
Q3 - Self-propagating malware infection (virus or worm)	No gap	$\mu0\text{-}3 = \mu1\text{-}3$
Q4 - Non-propagating malware infection (Trojan or rootkit)	Low	$\mu0\text{-}4 < \mu1\text{-}4$
Q5 - Denial of Service attacks	No gap	$\mu0\text{-}5 = \mu1\text{-}5$
Q6 - Theft or breach of proprietary or confidential information	Medium	$\mu0\text{-}6 < \mu1\text{-}6$
Q7 - System penetration by outsider	Low	$\mu0\text{-}7 < \mu1\text{-}7$
Q8 - Laptop theft	No gap	$\mu0\text{-}8 = \mu1\text{-}8$
Q9 - Insider abuse of Internet access, e-mail or computer resources	Low	$\mu0\text{-}9 < \mu1\text{-}9$
Q10 - Unauthorized access to information by insider	Medium	$\mu0\text{-}10 < \mu1\text{-}10$
Q11 - Website defacement	Low	$\mu0\text{-}11 < \mu1\text{-}11$

In accordance with Table 5-20 and Table 5-21, the presence of 'medium' and 'low' and 'no' gaps were highlighted, whilst the characteristics of a 'high' gap could not be found.

The risks that are perceived with the *'highest'* gap (a gap rated as *medium* [+/- 2 points] in accordance to our rating scales) are the ones from question 6 and 10, respectively the likelihood of the risks are:

- Theft or breach of proprietary or confidential information; &
- Unauthorized access to information by insider.

The risks that are perceived with a gap of '*low*' (rated [+/- 1 points] in accordance to our rating scales) are the ones from Question, 2, 4, 7, 9 and 11 respectively the likelihood of the risks:

- Virus attacks;

- Non-propagating malware infection (Trojan or rootkit);
- System penetration by outsider;
- Insider abuse of Internet access, e-mail or computer resources; &
- Website defacement.

The risks that are perceived as no different (*no gap* in accordance to our rating scales) are those from question 3, 5 and 8, respectively the likelihood of the risks:

- Self-propagating malware infection (virus or worm);
- Denial of Service attacks; &
- Laptop theft.

In addition, the research results can provide the direction of the gaps as shown in Table 5-21.

Question 2, 'virus attacks', was the only risk that was rated by the respondents with a lower score than the one obtained by the information security surveys ($\mu0\text{-}2 > \mu1\text{-}2$). All other risks were rated as having a higher likelihood than the one recorded by the surveys.

In accordance with the literature review (Section 2) those results demonstrate an under-estimation of the 'likelihood of virus attacks' whilst a general over-exaggeration of the likelihood of the other risks seem to be predominant. The only exceptions to these observations are question 3, 5 and 8 where the likelihood of those risks did not significantly differ from the baseline set by the surveys.

In other words, the likelihood of 'Virus attacks' (question 2) was perceived by the participating sample as lightly to moderately lower (under-estimation) than it was in the surveys. On the other hand, the likelihood of the risks of 'Non-propagating malware infection (Trojan or rootkit)', 'Theft or breach of proprietary or confidential information', 'System penetration by outsider', 'Insider abuse of Internet access, e-mail or computer resources', 'Unauthorised access to information by insider' and 'Website defacement' were all perceived by the participating sample as lightly to mildly higher (over-exaggeration) than it was in the surveys.

5.5. Association between self-rated knowledge and risk rating scores

Question 1 was concerned with the self-rating of the participant's information security knowledge. As seen in Section 3.6, those self-rating scores were captured using a wide six-point scale varying from 'No knowledge' (0) to 'Expert' (5).

On the other hand, questions 2 to 11 were used to highlight a measure of the *gap* between an individual's perceptions of risks against their ranked occurrences as shown by the information security surveys.

Social perceptions and representations have been speculated to play an important role in risk perceptions. Section 2 of this research reviewed some of the literature and broad concepts surrounding social influences and self-representations. To factor those elements of research into the data analysis the following hypothesis was formulated based on the assumptions reviewed in the literature review.

H_0: 'The association exists between self-assessed information security knowledge ratings and perceptions of risks.'

This section does not intend to analyse every trend and correlation occurring between every variable other than the ones specified above. In order to do this the *bivariate* output analysis was strictly limited to the analysis of those two variables.

Bivariate correlations

Bivariate is a statistical method used to measure the degree of correlation between two variables, thereby, determining their statistical independence and measure of association.

For the purpose of this research the following scope of the analysis was limited to the following two variables:

- Participant's self-rated knowledge score and
- Participant's answers to ten risk questions.

As the above two variables are categorical to each question associated to each risk, Jackson (2006) recommends the use of the *Spearman's rho (rho)* as the appropriate correlation method for those variables.

The *correlation coefficient*, which ranges from '-1' to '+1', is both a measure of the strength and the direction of the relationship. Furthermore, for the purpose of this research, a *one-tailed test* was used to determine the statistical significance of this *correlation coefficient.*

The expected direction of the association is an inferred strong positive relationship between expert knowledge and risk perception. The *one-tailed test*s is used to confirm or reject *null* hypothesis (H_0) and determine whether or not the correlation is statistically, significantly different from zero. If there is no significance, the value of the *one-tailed test* would not provide any conclusive information as it can be interpreted only in the context of the criterion demonstrating significance.

Caution must be observed with the analysis of the *correlation coefficient,* as correlation is not causation. By looking at the correlation between subjects that self-rated their information security knowledge (Q1) and the risk ratings (Q2-11), the data could show a positive, negative or no correlation; yet the researcher would be reluctant to claim that those subjects always perceive risks this way. Perception of a risk can be changed, the context may change, many variables can fluctuate with the data and therefore correlation cannot be causation.

Analysis of the association

Table 5-22 demonstrates the *bivariate* correlation between self-rated knowledge for the thirty-seven (N=37) respondents and their answers to the ten (10) risk questions.

Based on test results obtained using *Spearman's rho* (*rho*), at the required confidence level (p=0.05), the *null* hypothesis, namely that no association exists between self-assessed information security knowledge and perceptions of risk, could not be excluded.

A slightly positive relationship exists between question 1 and question 9, 10 and 11, however not statistically significant enough to have an impact on the correlation of those variables. Similarly, a slightly negative relationship exists between question 1 and questions 2, 3, 4, 5, 6, 7 and 8, however, not statistically significant enough to have an impact on the correlation of those variables.

Table 5-22 Results from one-tailed testing with Spearman's Rho

			Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
Spearman's rho	Q1	Correlation Coefficient	1.000	-.087	-.010	-.232	-.114	-.180	-.116	-.133	.079	.169	.000
		Sig. (1-tailed)	.	.305	.476	.084	.250	.143	.247	.217	.321	.158	.498
		N	37	37	37	37	37	37	37	37	37	37	37

**. Correlation is significant at the 0.01 level (1-tailed).

*. Correlation is significant at the 0.05 level (1-tailed).

The *one-tailed test* does not demonstrate any relevant information in this instance as it can only be interpreted if the effect meets the criterion for significance, which in this case it did not for the variables used.

Based on the rejection of the *null* hypothesis the data describes a situation in this study where no statistical proof could be established for a change in one variable (i.e.: self-rated knowledge) being associated with any particular change in another variable (i.e. risk rating).

Research Question: 'Are there risk perception differences between people that self-assessed their information security knowledge above or below average and the overall participating sample?'

In the absence of any argument to demonstrate a correlation between the two assessed variables, no association between them can be inferred, in the context of this research. The data do not establish a conclusive argument to demonstrate that people who self-assessed as 'above average' (or below average) rated those risks differently to other subjects in the sample.

In other words, knowing the value of the self-assessed knowledge of a subject does not provide this research with information about the way this subject rated the ten risk questions.

The sample population used in this research mostly self-rated their knowledge at a similar level (μ1-1=3.49; SD=0.989). Such reduced variability in the participant's answers could explain the observed outcome.

6. Conclusions

6.1. Discussion

The fundamental research question was: 'is there a gap between people's perception of information security risks and the survey results?'

The analysis of the results described in Section 5 indicates the existence of such a gap. This gap varied for different risks.

The risks that were perceived with the *'highest'* gap were associated to the likely occurrence of the 'theft or breach of proprietary or confidential information' and 'unauthorized access to information by insiders'. The risks that were perceived with the *'lowest'* gap were associated with the likelihood occurrence of 'virus attacks', 'non-propagating malware infection (Trojan or rootkit)', 'system penetration by outsiders', 'insider abuse of Internet access, e-mail or computer resources' and 'website defacement'. Only one of those risks ('viruses attacks') was perceived as occurring less frequently than the actual risk (according to the information security surveys). All other risks seemed to be over-exaggerated in the perception of the participants.

The risks that were perceived as no different were associated to the likely occurrence of 'self-propagating malware infection (virus or worm)', 'denial of Service attacks' and 'laptop theft'.

The risk perception theories previously reviewed state that people are subject to certain mental shortcuts that affect their perception of this risk when confronted with it. That perception influences the decision making process with regard to those risks.

The existence of those mental shortcuts (biases and heuristics) influenced the researcher's hypothesis to expect similar results to the ones presented in Table 2-2. The results supported the risk perception theories. An informal correlation between the highlighted gaps and the conventional wisdom about risks outlined by Table 2-2 implies that the risks concerned by perception differences are very much aligned with the risk perception theories described in the literature review of this research. People have a tendency to exaggerate risks that are:

- rare;
- talked about;
- intentional or man-made;

- affecting them personally;
- new and unfamiliar; &
- morally offensive.

Most of those attributes can be associated with the over-exaggeration of the risks which had identified gaps.

The risks with the highest perception gaps regularly seem to be the subject of recent news reports, for example: privacy breaches, use and misuse of personal data with social networking sites, loss of confidential and credit card information by large organisations). Psychological literature explains that the 'availability heuristic' is at play when people assess the likely occurrence of risks. This concept infers that data that is easily available (from memory, as recently seen in the media, etc.) are given more weight than other data. This reflects a description of most risks affected by perception gaps.

This research would require further data collection from the participating sample, the development of a new questionnaire using the same ten risks but tailored to capture information relevant to the availability theory (i.e.: 'Have you experienced a Denial of Service attack over the past 1 month? 3 months? 6 months?'). This could provide further data to enable a correlation between the overestimation in a risk rating and the availability of these risks. This additional research could confirm how sound those gaps are in accordance with risk perception theories.

In the context of information security, risk assessments are valuable to identify, analyse and evaluate risks. Security experts are a trusted resource for evaluating risk and apply their knowledge objectively. However, there is a general tendency to accept the analysis as outlined by those experts which could potentially lead to a false sense of security or wrongful prioritization of risks due to human biases. In recent years, research provided evidence of the subjective nature of risk that is inherent in human nature. If lay people exhibit biases towards risks, why would experts be exempted from those perceptions?

To capture the disparities between people's judgments and risk perceptions, this research asked all participants to self-assess their level of information security knowledge on a six-point scale. The data obtained from this question were then statistically analysed using correlation methods against the ratings of ten risk questions.

This research found no statistically significant evidence that the subjects with a self-rated information security knowledge of ‘above average’ rated risks any differently than the other participants in the sample. However, the absence of evidence is not evidence of absence. It would be wrong to infer that such correlation does not exist at all. The majority of the respondents rated their knowledge as ‘Good’, which in accordance with the descriptive statistics analysis depicts little variability in the types of answers given. This could explain the outcome. Similarly, the insufficiency of the data or the context used in this research could have been a limiting factor. Therefore, one outcome is that purposeful subject selection should be used in any future research to construct a sample with greater variability in terms of self-assessed knowledge. Such a sample may prove difficult to construct if further investigation shows a Hawthorne type effect (Jackson, 2006) in which subjects consistently over rate their state of knowledge about computer security.

6.2. Future research directions

The scientific rationality used in risk analysis is expressed through a rational enquiry and procedure set which leverages mathematical values and metrics to typically evaluate risks as well as their likelihood and impact.

Such analysis significantly contributes to decision-making; it is therefore important to ensure those values, as well as their limitations, are correct.

Risk perception theories account for those limitations by acknowledging that people judge risks subjectively based on their characteristics and severity. This in turns leads to the overestimation of certain risks and the underestimation of others.

The literature of risk perceptions covers many of the elements at stake into two major families: the psychometric paradigm and cultural theory.

Both theories seek to explain why people estimate the level of danger of specific risks differently. Thus, understanding the subjectivity and the existence of risk perceptions as explained through those theories should be highly beneficial to a risk analysis process.

This research has just started to explore the relevant literature in psychology and risk. There is a great deal more to research and learn from the correlations of risk perception, behavioural sciences, psychology of risk and decision making related fields of study.

However, this research identified the existence of risk perception gaps between the participating sample and the results from the selected information security surveys.

The next logical step would be to conduct further experiments and analysis of the identified gaps with a deeper psychological analysis of the biases and heuristics affecting the participants and relate it more specifically to the analysis of risks.

Another research direction would be the validation and correlation of the highlighted perception gaps using other sources of data. This research leveraged results from the AusCERT and CSI/FBI survey data for 2005 to 2007. However, to ensure that the gaps can be reproduced other data sources would add a validation layer to the initial results and ensure results are scientifically sound.

A long-term goal of this research is the integration of its findings into a conceptual risk assessment model or framework. This model could bring a closer focus to factoring risk perceptions that occur when humans analyse risks and make decisions based on those perceptions.

This would require the establishment of a context that takes into account both realistic and subjective aspects of risks.

References

AISA. (2007). Australian Information Security Association. Retrieved February 28, 2008, from http://www.aisa.org.au/

Alberts, C. J., & Dorofee, A. J. (2001). *Octave Criteria, version 2.0.* Pittsburgh: Carnegie Mellon Software Engineering Institute.

Ames, M. (2007). *Risk Management in Context.* Paper presented at the Oceania CACS Conference 2007, New Zeland. Retrieved September 2007.

AS/NZS:4360:2004. (2004). Risk Management Guidelines. In S. Global (Ed.) (third ed.): SAI Global.

AusCERT. (2005). 2005 Australian Computer Crime and Security Survey. Retrieved September 15, 2007, from http://www.auscert.org.au/images/ACCSS2005.pdf

AusCERT. (2006). 2006 Australian Computer Crime and Security Survey. Retrieved September 15, 2007, from http://www.auscert.org.au/images/ACCSS2006.pdf

Bailey, R. (2006). Don't Be Terrorized. Retrieved September 29, 2007, from http://www.reason.com/news/show/36765.html

Brenner, B. (2006). Has CSI/FBI survey jumped the shark? Retrieved September 15, 2007, from http://searchsecurity.techtarget.com/columnItem/0,294698,sid14_gci1202328,00.html

Burke, R. J., & Greenglass, E. R. (2004). *International Journal of Stress Management, 7*(1), 49-59.

Byrne, D. (2003). *"Irrational Fear or Legitimate Concerns". Risk Perception in Perspective.* Paper presented at the European Commissioner for Health and Consumer Protection, Brussels. Retrieved 2007.

Cooper, W. H. (1981). Ubiquitous halo. *Psychological Bulletin, 90*, 218-224.

CSI/FBI. (2006). Computer Crime and Security Survey. Retrieved September 15, 2007, from http://www.gocsi.com/forms/fbi/csi_fbi_survey.jhtml

CSI/FBI. (2007). Computer Crime and Security Survey. Retrieved September 15, 2007, from http://www.gocsi.com/forms/fbi/csi_fbi_survey.jhtml

Deloitte. (2006). 2006 Global Security Survey. Retrieved September 15, 2007, from http://www.deloitte.com/dtt/cda/doc/content/dtt_fsi_2006%20Global%20Security%20Survey_2006-06-13.pdf

Dijksterhuis, A., Bos, M. W., Nordgren, L. F., & Baaren, R. B. v. (2006). On Making the Right Choice: The Deliberation-Without-Attention Effect. *Science, 311*, 1005-1007.

Douglas, M. (1994). Risk and Blame: Essays in Cultural Theory. *The British Journal of Sociology, 45*(1), 143-144.

ENISA. (2006a). *Risk Assessment.* Retrieved September 22, 2007. Retrieved September 18, 2007, from http://www.enisa.europa.eu/rmra/rm_process_02.html

ENISA. (2006b). Risk Management: Implementation principles and Inventories for Risk Management/Risk Assessment methods and tools. Retrieved October 14, 2007, from http://www.enisa.europa.eu/rmra/files/D1_Inventory_of_Methods_Risk_Management_Final.pdf

Finucane, M., Slovic, P., Mertz, C. K., Flynn, J., & Satterfield, T. A. (2000). Gender, Race and Perceived Risk: The 'White Male' Effect. *Health, Risk & Society, 2*(2), 159-172.

GAO. (1998). Executive Guide: Information Security Management. Retrieved September 16, 2007, from http://www.gao.gov/archive/1998/ai98068.pdf.

Guillot, A., & Kennedy, S. (2007). *Information Security Surveys: A Review of the Methodologies, the Critics and a Pragmatic Approach to their Purposes and Usage.* Paper presented at the 5th Australian Information Security Management Conference, Edith Cowan University, Perth, WA. Retrieved.

Hasehuhn, M. P., & Mellers, B. A. (2005). Emotions and cooperation in economic games. *Cognitive Brain Research, 23*, 24-33.

Jackson, S. L. (2006). Research Methods and Statistics: A Critical Thinking Approach (2 ed.). Belmont: Thomson Wadsworth.

Joffe, H. (2003). Risk: From perception to social representation. British Journal of Psychology, 42(1), 55-73.

Judd, C. M., Smith, E. R., & Kidder, L. H. (1991). *Research methods in Socail Relations* (Sixth ed.). Orlando, Florida: Ted Buchholz.

Kahneman, D. (2003). A Perspective on Judgment and Choice: Mapping Bounded Rationality. *American Psychologist, 58*(9), 497-720.

Kahneman, D., & Tversky, A. (1984). Choices, Valus, and Frames. *American Psychologist, 39*(4), 341-350.

Krech, D., & Crutchfield, R. S. (1948). *Theory and problems of social psychology.* New York: McGraw-Hill.

Lerner, J. S., Small, D. A., & Loewenstein, G. (2004). Heart strings and purse strings: Carryover effects of emotions on economic decisions. *Psychological Science, 15*(5), 337-341.

Marris, C., Langford, I. H., & O'Riordan, T. (1998). A Quantitative Test of the Cultural Theory of Risk Perceptions: Comparison with the Psychometric Paradigm. *Risk Analysis, 18*(5), 635-646.

Octave. (2003). Information Security Risk Evaluation. Retrieved Semptember 28, 2007, from http://www.cert.org/octave/

Osman, A., Gutierrez, P. M., Barrios, F. X., Kopper, B. A., & Chiros, C. E. (1998). The Social Phobia and Social Interaction Anxiety Scales: Evaluation of Psychometric Properties. *Journal of Psychopathobgy and Behavioral Assessment, 20*(3), 249-265.

Schneier, B. (2007). The Psychology of Security. Retrieved September 27, 2007, from http://www.schneier.com/essay-155.pdf

SCISSEC. (2007). Research Group. Retrieved September 25, 2007, from http://scissec.scis.ecu.edu.au/wordpress/

Shrauger, J. S., & Osberg, T. M. (1981). the relative accuracy of self-predictions and judgments by othersin psychological assessment *Psychological Bulletin, 90,* 322-351.

Skjong, R., & Wentworth, B. H. (2001). *Expert judgment and risk perception.* . Paper presented at the SOPE-2001: Eleventh (2001) International Offshore and Polar Engineering Conference, Norway. Retrieved October 8 2007.

Slovic, P. (1987). Perception of risk. *Science, 236,* 280-285.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1980). "Facts and fears: understanding perceived risk". In R. C. Schwing & W. A. Albers (Eds.), *Societal Risk Assessment. How Safe is Safe Enough?* (pp. 181-216). London: Pienum.

Slovic, P., Fischhoff, B., Lichtenstein, S., & Roe, F. J. C. (1981). Perceived Risk: Psychological Factors and Social Implications. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 376*(1764), 17-34.

SPSS [Computer software]. (2008). SPSS Inc., Chicago: Illinois

Tan, A. (2006). The weakest link in the security chain? You. Retrieved September 29, 2007, from http://software.silicon.com/security/0,39024655,39158023,00.htm

Theronsolutions. (2006). The CSI/FBI 2006 Survey Considered Irrelevant. Retrieved September 15, 2007, from http://theron.com.my/blog/2006/07/17/the-csifbi-2006-survey-considered-irrelevant/

Thompson, M., Ellis, R., & Wildavsky, A. (1990). *Cultural Theory*. Boulder: Westview Press.

Trochim, W. (2006a). Research Methods Knowledge Base: Likert Scaling. Retrieved September 30, 2007, from http://www.socialresearchmethods.net/kb/scallik.php

Trochim, W. (2006b). Research Methods Knowledge Base: Nonprobability Sampling. Retrieved September 30, 2007, from http://www.socialresearchmethods.net/kb/sampnon.php

Trochim, W. (2006c). Research Methods Knowledge Base: Positivsm & Post-Positivsm. Retrieved September 30, 2007, from http://www.socialresearchmethods.net/kb/positvsm.php

Trochim, W. (2006d). Research Methods Knowledge Base: Types of design. Retrieved September 30, 2007, from http://www.socialresearchmethods.net/kb/destypes.php

Tucker, S. (2007). Distance Education: Better, Worse, Or As Good As Traditional Education? Retrieved October 2007, 2007, from http://www.westga.edu/%7Edistance/ojdla/winter44/tucker44.html

Tversky, A. (1974). Assessing Uncertainty. *Journal of the Royal Statistical Society, 36*(2), 48-159.

UCCASS. (2004). Unit Command climate assessment and Survey System. Retrieved December 8, 2007, from http://www.bigredspark.com/survey.html

WAISSIG. (2007). West Australian Information Security Special Interest Group Retrieved November 8, 2007, from http://www.waissig.org/index.html

Walsh, C. (2006). CSI/FBI Survey considered harmful. Retrieved September 15, 2007, from http://www.emergentchaos.com/archives/2006/07/csifbi_survey_considered.html

Whitman, M. E., & Herbert, M. J. (2005). Risk Assessment. In *Principles of Information Security*. Canada: Thomson.

Winkler, I. (2006). Opinion: Investigating the FBI's 'invalid' security survey. Retrieved September 15, 2007, from http://searchsecurity.techtarget.com/columnItem/0,294698,sid14_gci1159992,00.html

Appendix

A. Questionnaire

1. How would you rate your Information Security knowledge on the following scale?

□ None

□ Some

□ Average

□ Good

□ Expert

PLEASE CONSIDER THE FOLLOWING BEFORE ANSWERING QUESTIONS 2 TO 11

Based on your own understanding and/or perception of the given risks, please rate the following:

2. In your opinion, rate the likelihood of virus attacks?

□ Very Low

□ Low

□ Medium

□ High

□ Very High

3. In your opinion, rate the likelihood of self-propagating malware infection (virus or worm)?

□ Very Low

□ Low

□ Medium

□ High

□ Very High

4. In your opinion, rate the likelihood of non-self propagating malware infection (Trojan or rootkit)?

□ Very Low

□ Low

□ Medium

□ High

□ Very High

5. In your opinion, rate the likelihood of Denial of Service attacks?

□ Very Low

□ Low

□ Medium

□ High

□ Very High

6. In your opinion, rate the likelihood of Theft or breach of proprietary or confidential information?

□ Very Low

□ Low

□ Medium

□ High

□ Very High

7. In your opinion, rate the likelihood of System penetration by outsider?

□ Very Low

□ Low

□ Medium

□ High

□ Very High

8. In your opinion, rate the likelihood of Laptop theft?

□ Very Low

□ Low

□ Medium

□ High

□ Very High

9. In your opinion, rate the likelihood of Insider abuse of Internet access, e-mail or computer resources?

□ Very Low

□ Low

□ Medium

□ High

□ Very High

10. In your opinion, rate the likelihood of unauthorised access to information by insider?

□ Very Low

□ Low

□ Medium

□ High

□ Very High

11. In your opinion, rate the likelihood of Website defacement?

□ Very Low

□ Low

□ Medium

□ High

□ Very High

B. Information security surveys selection criteria

Information security survey selection methods

Numerous security professionals use the statistics from these surveys for conferences, training and awareness campaigns (Theronsolutions, 2006)

The abundance of surveys in the information security area made it necessary to reduce their number. To overcome this issue a series of selection criteria were defined and applied to the different surveys. The following criteria were used in selecting the surveys used in this study:

- The survey must be directed specifically at information security risk;
- Outline statistical measurements of the security items measured:
- Some surveys are advisory only and are limited to concluding arguments in regards to previously analysed measurements.
- For the purpose of this paper, the survey must present the statistical information.
- Include a justification of the methodology and clearly state the nature of the respondents (i.e.: IT managers);
- Have a minimum of 3 years existence and include in the presentation of the results a comparison to the previous years in order to assess the trends over the recent years;
- Be selected from sources representative of current security professionals including, but not limited to, vendors and industry leaders, independently mandated surveys and governmental reports.
- The surveys must be freely available to the community.
- Note: Both international and country specific surveys were considered.

Summary and methodologies

The scope and depth of the information presented by each survey varies, rendering a comparison harder to make. The methodologies employed by the participants show similarities in the procedures taken to design collect and review the questionnaires.

In order to maximise the accuracy and fairness of the methodology review the information collection was limited to the following criteria:

- Audience surveyed;
- Dates and timelines associated with the surveys;
- Any available information regarding the design, collection and analysis of the information (i.e.: by whom, collection method, parties involved);
- Response-rate.

It is important to note that any information transcribed below is directly quoted from the associated survey.

AusCERT: 2006 Australian Computer Crime and Security Survey:

The Australian High Tech Crime Centre (AHTCC) the Australian Federal Police (AFP) the Police from NSW, QLD, SA, Tasmania, VIC, WA, Northern Territory, and AusCERT has collaborated to produce this survey. The survey was funded by the Australian government's Attorney-General's Department, ACNielsen, a market research and Information Company was engaged to assist with the preparation and conduct of the survey.

The survey was adapted from the 2006 CSI/FBI survey, providing the opportunity to compare Australian findings with the United States in some areas (AusCERT, 2006).

- The survey was deployed May 22, 2006
- Respondent answers cover the 12 months period before January 2006.
- Business reply-paid envelopes were sent to 2024 IT managers or their equivalents from a range of Australian public and private sector organizations. Those organizations were invited to complete the survey on-line or return the paper questionnaire via the reply-paid envelope. Responses were also sought from a number of private and public sector industry groups, including the Trusted Information Sharing Network (TISN) whose members were invited to complete the survey via the secure web site.
- Yielding 389 respondents (17% of response rate).
- All responses were anonymous.

- Readers were warned that the format of the survey changed to increase the survey sample size; this should consequently be considered when assessing the respondent percentages against previous years.
- This survey is adapted from the CSI/FBI survey (AusCERT, 2006).

CSI/FBI: Computer Crime and Security Survey

This survey is conducted by the CSI with the participation of the San Francisco Federal Bureau of Investigation's (CSI/FBI) Computer Intrusion Squad.

Its aim is to raise the level of security awareness, as well as help determine the scope of computer incidents in the United States (U.S.) (CSI/FBI, 2006).

- The survey was deployed early January 2006.
- Respondent answers cover the year of 2006.
- Hardcopy (fist-class mailing) and e-mail versions of the survey were distributed to 5000 information security practitioners in U.S. corporations, government agencies, financial institutions, medical institutions, and universities.
- Yielding 616 respondents (12.3% of response rate).
- All responses were administered anonymously.

Table B- 1 Summary of the Sponsors, Managers, Participants and Methodologies for the selected surveys

Surveys	AusCERT 2006	CSI/FBI 2006
Country surveyed	Australia	United States of America
Profile	Independent Non-profit	Public association (CSI/FBI) / Governmental (CSI/FBI)
Sponsors or cooperation by	Managed by: ACNielsen Sponsored by: AFP and all police forces, AHTCC, Attorney-General's Department	Managed by: Computer Security Institute (CSI/FBI) and the computer intrusion squad from San Francisco (CSI/FBI)
Frequency	Yearly	Yearly
Participants	IT managers or equivalent TISN members A number of private and public sector industry groups	CSO readers
Sample size	2024 (389 responded)	15 000 (434 responded)
Response rate	17%	2.89%
Survey method	Mail hard-copy Web-site	E-mail

C. Information letter and informed consent

Information letter

To whom it may concern,

I would like to extend an invitation for you to participate in an on-line survey as part of research being conducted on security risks. This research investigates the existence of perceptions in the context of information security risks.

The highlights of those perceptions and their nature have a significant impact on the risk assessment by introducing social, emotional and human considerations in its process as well as improve the quantification of information security risks by targeting risks more specifically. Security professionals can leverage those findings to make informed decisions in regards to those risks, thereby improving the management of those risks.

The on-line survey will be available for a two weeks period over a secure URL. The participants' answers will be anonymous and directly input in a password locked database, only accessible to the researcher. The data will be stored on removable medium in a secure encrypted format and will not be used for any purposes other than its intended use within this research; all collected data will be preserved for a period of not less than 5 years.

We value your opinions and appreciate your participation. Please use care when completing this survey as all data are collected anonymously, we highly encourage you to provide accurate answers.

If you wish to contact me to further discuss the implications of my research, please contact me by e-mail aguillot@student.ecu.edu.au.

Yours sincerely;

Alexis Guillot

B.IT Honours Student; Edith Cowan University

Informed consent

I confirm that I have received, read and understood the information provided in the Information Letter explaining this research study. I have been given the opportunity to ask questions and any questions have been answered to my satisfaction. I am aware that if I have any additional questions I can contact the researcher. I understand that participation in this research project will involve answering eleven opinion-based questions on the subject of information security risks. I understand that my answers are anonymous and will be kept confidential that they will be used only for the purposes of this research project only. I am aware that I am free to withdraw from further participation at any time, without explanation or penalty.

By taking this survey, you agree to freely participate in this project and are over 18 years of age.

D. Ethics approval letter

Edith Cowan University **Memo**

Faculty of Computing, Health and Science

Human Ethics Subcommittee

TO: Tamara Harold, Admin. Officer, Higher Degrees

FROM: Angus Stewart, Chair, Faculty Human Ethics Subcommittee

SUBJECT: Human Ethics Clearance Application/s

DATE:29th January, 2008

Dear Tammie,

The following ethics application by

Alexis Guillot	Is there a gap between people's perception of information security risks and the findings of published researches of those risks.

is approved (category 2), subject to the following:

All collected data should be preserved for a period of not less than 5 years.

Best wishes,

Angus.

E. Survey results: exported data

Data were retrieved using the following SQL query. Those data are shown in Table E-1 below.

```
SELECT r.sequence, q.Question, a.value FROM `alexisresults` r, `alexisquestions` q,
`alexisanswer_values` a WHERE r.`sid` =29 AND r.qid = q.qid AND a.avid = raved
```

The first column represents a generic number associated with each respondent. The numbering sequence in the following table does not represent anything in particular to the researcher and should be interpreted as follows: the first sets of answers for respondent number one is labelled 1, the second sets of answers by respondent number two is labelled '2' and so on for the whole 36 participants.

Table E- 1 Results sorted per Question

Respondent Nb#	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
1	Good	Medium	Low	Low	Low	Medium	Low	High	Medium	High	Low
2	Average	High	High	High	Low	Medium	Medium	High	High	High	Low
3	Some	High	High	High	Low	High	Low	High	Very High	Medium	Low
4	Good	High	High	Very High	Low	Low	Low	High	Very High	High	Low
5	Expert	Very High	Low	Medium	High	Very High	Medium	Very High	High	Very High	Medium
6	Good	Low	Medium	Medium	Low	High	Low	High	Very High	Very High	Low
7	Expert	High	High	High	Low	Very Low	Low	High	High	Very High	Low
8	Good	Medium	Low	Low	Low	Low	Low	High	Very High	Low	Low
9	Good	High	High	Medium	Medium	High	Medium	Medium	High	High	Low

10	Good	High	High	High	Low	Medium	Medium	Medium	Medium	High	Medium
11	Average	High	High	High	Medium	High	Medium	High	High	High	Medium
12	Good	Low	Low	Low	Low	Very Low	Medium	High	Medium	Medium	Low
13	Average	Medium	Low	High	High	Very High	Medium	Medium	Very High	High	Low
14	None	High	Medium	Medium	Medium	High	High	Very High	Medium	Medium	Low
15	Average	High	Medium	High	Medium	Medium	High	Medium	Very High	High	Medium
16	Good	Medium	High	Low	Very Low	Very Low	Very Low	Low	High	Very Low	Very Low
17	Some	Very High	Very High	High	Medium	High	Medium	Very High	High	High	Medium
18	Good	High	Medium	High	Medium	High	Medium	Very High	Very High	Low	Low
19	Average	High	Medium	High	Medium	High	Low	Very	Very	High	Very

								High	High		Low
20	Good	Medium	High	High	Low	High	Low	Low	Very High	High	Low
21	Good	Medium	Medium	Medium	Medium	Low	Medium	Low	Low	Medium	High
22	Good	Medium	High	High	High	High	Low	Very High	Very High	High	Medium
23	Some	Low	Medium	Low	High	Very Low	Low	High	Very High	High	Medium
24	Good	High	High	Medium	Medium	Medium	Low	Medium	Very High	Medium	Low
25	Average	Low	Medium	Low	High	High	Medium	High	Very High	High	High
26	Good	Medium	High	Medium	Very High	High	Medium	Low	High	Medium	Very Low
27	Some	Medium	Medium	High	Low	Medium	Medium	Very High	Medium	Medium	Medium
28	Average	Low	Low	Low	Low	Medium	Low	High	Very High	Low	Low

29	Good	Low	Low	Low	Low	High	Medium	High	Very High	Medium	Low
30	Average	High	High	Medium	Low	Medium	Medium	Low	High	High	Very Low
31	Good	Medium	Medium	Medium	Very Low	High	High	Medium	Medium	High	High
32	Good	Very High	Very High	High	Low	Medium	Low	High	Very High	Medium	Medium
33	Average	High	High	High	Medium	High	Low	Low	Medium	Medium	Low
34	Expert	Medium	High	Low	Medium	Medium	High	High	Very High	High	Medium
35	Average	Very High	High	High	High	Very High	High	High	Medium	Low	High
36	Good	High	Very High	Medium	High	Medium	Low	High	Very High	High	Medium
37	Good	Medium	Low	Low	High	High	High	Very High	Very High	Medium	High