ABSTRACT

USER EVALUATION OF THE PERFORMANCE OF INFORMATION SYSTEMS

by
Edward Mahinda

Information technologies (IT) are considered the primary survival factor for many organizations and the most critical success factor in businesses today. To justify the necessary investment in IT, user evaluation of information systems' performance in organizations is a key consideration. This research investigated a comprehensive and convenient means for end users to assess this performance.

Among the existing theories and models on the evaluation of information system performance based on intrinsic technological properties, the Web of System Performance (WOSP) model provides the most comprehensive basis for information system evaluation, and therefore merited further investigation. The research question was how well the eight WOSP performance criteria, namely functionality, usability, flexibility, reliability, security, extendibility, connectivity, and privacy, applied in the context of an individual evaluating one or more information systems for use by an organization.

For this, it was important to show that, while these performance criteria were abstract concepts, they can be established and identified clearly, in a manner that is valid in the sense of the meaning and that users would consider important. Illustrative statements for each of the eight criteria were therefore obtained, which users were asked to evaluate.

Next, it was necessary to show that users prefer the choice of the eight WOSP criteria to the current dominant instrument for evaluation when evaluating software. This was done using a preference questionnaire where subjects rated both the WOSP model and an alternative means of evaluation along various dimensions, the results being compared by statistical analysis

Finally, it was necessary to show that users rate at least three of the WOSP criteria as being important for evaluating information systems. For this, conjoint analysis was used. A browser was selected as the experimental software for this research.

The results showed that users found illustrative statements clear, valid and important for the evaluation of browsers. They also preferred using the WOSP model for the evaluation of browsers over TAM, the current dominant model. Finally, while users attached different levels of importance to the various performance criteria for the selection of browsers, five of the criteria were important to a significant degree.

USER EVALUATION OF THE PERFORMANCE OF INFORMATION SYSTEMS

by
Edward Mahinda

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Information Systems

Department of Information Systems

January 2008

APPROVAL PAGE

USER EVALUATION OF THE PERFORMANCE OF INFORMATION SYSTEMS

Edward Mahinda

12/12/07

Dr. Brian Whitworth, Dissertation Advisor — Date
Senior Lecturer, Institute of Information and Mathematical Sciences
Massey University, Albany, Auckland, New Zealand
Assistant Professor of Information Systems (Until January 2006)
College of Computing Sciences, NJIT

11/5/07

Dr. George Widmeyer, Committee Member — Date
Associate Professor of Information Systems
College of Computing Sciences, NJIT

11/06/07

Dr. Cheickna Sylla, Committee Member — Date
Associate Professor of Management
School of Management, NJIT

11/06/07

Dr. Jerry Fjermestad, Committee Member — Date
Associate Professor of Management
School of Management,
Associate Professor of Information Systems
College of Computing Sciences, NJIT

11/20/07

Dr. Omar El Sawy, Committee Member — Date
Professor of Information Systems
University of Southern California

BIOGRAPHICAL SKETCH

Author: Edward Mahinda

Degree: Doctor of Philosophy

Date: January 2008

Undergraduate and Graduate Education:

- Doctor of Philosophy in Information Systems,
  New Jersey Institute of Technology, Newark, NJ, 2008

- Masters of Business Administratration,
  Cornell University, 2000

- Masters in Engineering (Civil),
  Cornell University, 1999

- Bachelor of Science in Civil Engineering
  University of Nairobi, 1980

Presentations and Publications:

Whitworth, B., Banuls, V., Cheickna, S., and Mahinda, E. (2007). Expanding the criteria for evaluating social-technical software. *(Accepted for publication IEEE SMC).*

Whitworth, B., Fjermestad, J. and Mahinda, E. (2006). The Web of System Performance: A multi-goal model of information system performance. *Communications of the ACM*, May, 49(5), 93-99.

Whitworth, B., Cheickna, S., and Mahinda, E.(2006). Assessing emergent business IT using the web of system performance. *The Fifth Wuhan International Conference on E-Business, Wuhan, P.R. China*, 1-15.

Mahinda, E and Whitworth, B.(2006). Using The WOSP Model To Improve End-User Productivity Of Information Systems. *International Conference on Business IT (BIZIT 2006): Collaborating with ICT Innovations for Business Survival.*

Whitworth, B., Banuls, V., Cheickna, S., and Mahinda, E. (2005). An Expanded View of Technology Acceptance. *Journal of AIS Sponsored Theory Development Workshop, Las Vegas*, 1-19.

Mahinda, E. and Whitworth, B. (2005). The Web of System Performance: Extending the TAM Model. *Americas Conference on Information Systems, Omaha, Nebraska, USA* ,367-374.

Patten, K., Whitworth, B., Fjermestad, J., and Mahinda, E. (2005). Leading IT Flexibility: Anticipation, Agility and Adaptability. *Americas Conference on Information Systems, Omaha, Nebraska, USA*, 2787-2792.

Mahinda, E. and Whitworth, B. (2004). Evaluating Flexibility and Reliability in Emergency Response Information Systems. *International Workshop on Information Systems for Crisis Response and Management, Tilburg, The Netherlands,* 1-6.

To my father, Godfrey Wanjohi Mahinda.
For his unwavering support, and an unshakeable faith in the power of knowledge
An enduring source of inspiration

ACKNOWLEDGMENT

I would like to put on record my deepest appreciation to Dr. Brian Whitworth, my research supervisor and dissertation committee chair. He was an invaluable source of ideas, insights, comments and encouragement, and provided immeasurable support that was above and beyond any call of supervisory duties, especially in many times of difficulty. I am very grateful to the members of my dissertation committee, Dr. Cheikna Sylla, Dr. Jerry Fjermestad, Dr. George Widmeyer, and Dr. Omar El Sawy who were very active indeed, and who by dint of eminence in their respective specialty, gave me extremely useful and timely guidance, critical for the successful outcome of my endeavors. In this regard, special thanks go to Dr. El Sawy for accepting to be an external member of my committee, and who very willingly put his immense experience and knowledge, both widely acknowledged and distinguished, at my disposal.

I owe special thanks to the Office of Graduate Students which not only funded a major part of my studies, but was also very sympathetic and helpful in assisting me overcome the various unexpected challenges I encountered in the course of my studies.

I would also like to thank the Chair of the Information Systems Department, Dr. Michael Bieber, who, together with the staff in his office, has been very accommodating of me; as well as faculty, staff and colleagues in the department who were helpful in ways too diverse to mention.

Finally, but certainly not least, I am most grateful to my friends and family who have very helpful in innumerable ways, and unquestioning of me in my long and eventful journey.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Significance of Evaluating Performance of Information Systems

The use of information technologies, including computers, in today's organizations has increased dramatically in recent years. A major reason for this is that Information Technology has for some time now been considered the primary survival factor for many organizations today (Sylla and Wen 2002). By some estimates, over the last two decades, approximately 50% of all new capital investment in organizations has been in information technology (Westland and Clark 2000). As far back as in the late '80s, 0.5 trillion dollars per year was already being spent on information systems. Further, firms were investing between 1.5% and 3.0% of their annual revenue in IT (Ian 1989), while approximately one third of the annual capital investments by US corporations was in IT. By other estimates, the total worldwide expenditure on information technology exceeded one trillion US dollars per annum in 2001, with approximately a 10% annual compounded growth rate (Seddon, Graeser et al. 2002). These estimates give an idea of the significance of IT in business and industry today.

In fact, it is clear that in recent years, IT, more than any other technology has been considered as the most critical success factor of business organizations in the increasingly global competitive environment. The benefits that can be derived from IT investment give ample reason for the interest. The benefits can broadly be classified into four categories with the following purposes (Sylla and Wen 2002):

- Increasing productivity and the performance of operating processes.
- Facilitating support for management.
- Gaining competitive advantage.
- Providing a good framework for the restructuring or transformation of business.

Increased productivity: The use of IT to substitute for human effort in complementing or automating tasks and business processes impacts on system effectiveness and efficiency. Possible benefits include reduced labor, process cycle times, and communication time; increased timeliness and data accessibility; and increased income from product/service quality improvements (Farbey, Land et al. 1993; Applegate, McFarlan et al. 1996).

Management support: Benefits related to management support are seen in the form of reduced decision times and improved decision quality, as well as improved communications through office automation and the standardization of work processes. Other benefits include increased flexibility, better control of business processes, improved customer services, and a more effective use of employees. These and other forms of managerial support by IT are extensively documented in standard texts on IT investment such as (Farbey, Land et al. 1993; Applegate, McFarlan et al. 1996).

Competitive advantage and business transformation: IT may be used to create competitive inequality to gain competitive advantage (Sylla and Wen 2002). This can lead to the following advantages for a firm:

- Increased market share
- Improved operating margins relative to competitors
- Differentiation of products and services by unique features

Business restructuring and transformation: In addition, IT has been used to leverage a firm's core competencies and to help integrate its resources and strategies (Sethi and King 1994). Recently, IT has become key to restructuring business processes under what is termed business process reengineering (BPR). Other IT-enabled transformations include the flattening of the organizational structure, the creation of virtual value chains among the business units of an organization, and the creation of "borderless" virtual firms. In these illustrations, IT enhances the firm's internal as well as external efficiency (Farbey, Land et al. 1993; Turban, Lee et al. 2000).

1.2 Description of Information Systems

A system can be defined as “a purposeful collection of interrelated components that work together to achieve some objective” (Sommerville 2004). In the specific case of information systems, the stimulus for system development typically is rooted in an organizational problem or opportunity (Browne and Rogich 2001). Systems that include software can be divided into two broad categories (Sommerville 2004):

Technical computer-based systems: These are systems that include hardware and software components but exclude procedures and processes. Examples include television, mobile phones, and most personal computer software. These systems are used by individuals and organizations for specific purposes. However, knowledge of these purposes is not part of the system.

Socio-technical systems: These include one or more technical systems, but also include knowledge of how the system should be used to achieve some broader objective. These systems have defined operational processes, and include people, who are the operators, as inherent parts of the system. Also, socio-technical systems are governed by organizational policies and rules, and may be affected by external constraints such as national rules and regulatory policies.

Some of the essential characteristics of socio-technical systems include the following:

1. They have emergent properties that are system-wide, rather than being particular to specific parts of the system. The emergent properties depend both on the system components and the relationships between them. Such properties can be evaluated only after the entire system has been assembled.

2. They are frequently nondeterministic. Thus, they may not always behave in the same way. For instance, when presented with a given a particular input, they may not always return the same output. One reason for this is that the system's behavior is dependent on human operators who do not always behave in the same way. Moreover, use of the system may create new relationships between the system components, leading to a change in its emergent behavior.

3. The extent to which the system supports organizational objectives depends not only on the system, but also on the stability of these objectives, the relationships and conflicts between them, and the interpretation of the objectives by the organization's staff. Thus reinterpretation of an objective that a system was successfully designed to support may result in the system being viewed as a failure.

Given the inherent flexibility of software, it is often used to rectify unexpected situations or problems in a system. For example, it is very common to opt to enhance the software capabilities of a system without increasing the hardware costs.

Of particular interest to information system research are socio-technical systems that include hardware and software, have defined operational processes, and offer an interface, implemented in software, to human users.

1.3 The Need for Evaluation of Information Systems

All the touted benefits of IT notwithstanding, many CEOs are unconvinced as to whether their investments in IT are worthwhile, or whether a particular technology is the most suitable for their business needs (Bidgoli 1996). Despite great efforts by organizations, analysts, and users, a majority of systems are either abandoned before they are completed or, even though completed, do not meet user requirements, all resulting in an estimated cost to organizations of more than $100 billion annually (Fischoff 1989; Standish 1996). Moreover, the kind of aforementioned investment notwithstanding, organizations today have fewer financial resources available for information technology than they previously did (Rivard, Poirier et al. 1997). This, in turn, has led to an increasing desire by organizations to control their spending on activities related to information systems, including end-user computing.

Thus, a key objective of much of the research done in IS has to do with assessing the value of the information technology in organizations and understanding the determinants of that value. The goal of such research is to help firms deploy and manage their IT resources better so as to enhance overall effectiveness (Taylor and Todd 1995). Furthermore, such research is critical in providing information that senior executives need to justify the huge investments in computers and related technology (Brynjolfsson 1993).

1.4 Alternative Approaches for Evaluating Information Systems

The evaluation of information systems can be done at the completion of various stages of the systems' development life cycle. Doing evaluation prior to undertaking the development of the systems is referred to as a feasibility assessment. When evaluation is done at the end of the next stage, which is the specification of requirements and logical design, it is referred to as specification and design review and approvals. Further evaluation is done at the end of physical design, coding or testing, which is termed acceptance testing and management review. Next, evaluations may be performed just after installation. This is referred to as post implementation review. Finally, evaluation of the system after it has been in place for a while is referred to as systems operations post installation review (Hamilton 1981). All of these different kinds of evaluations can be summarized and classified into two distinct categories:

- Formative evaluation gives information that is used during development to help improve the product under development, and is utilized by those engaged in the development process.

- Summative evaluation, on the other hand, is done after the completion of development. It furnishes information regarding the product's effectiveness to decision makers who are interested in its adoptability (Kumar 1990).

This research focuses on the assessment of information systems by users. It therefore concerns itself with summative evaluation.

There have been several approaches to the assessment and evaluation of information systems during use. One has been a macroeconomic approach to the whole question. This approach seeks to assess the value of IT based on the performance of the national economy. A major problem with it, however, is the fact that very many factors affect the national productivity, and so it is difficult to isolate the impact of IT. Another limitation has been the reliability of the data for this kind of analysis. Moreover, a significant number of sectors are not consistently defined over time (Brynjolfsson 1993). And even at the industry level, the correct picture is difficult to discern in some cases where better performing firms force their competitors out of the market, so that the productivity in that industry averages out (Brynjolfsson 1998).

A second approach has been at the level of the firm, assessing the relationship between IT expenditure and a firm's performance. Even at the firm level, it may be difficult to isolate the impact of IT alone, when for example investment in IT must be accompanied with other investments or changes in order to bear fruit. For example, firms might need to restructure the organization into a more decentralized structure and/or empower their employees more in order to more fully realize the benefits of IT (Brynjolfsson 1998).

However, in order that technologies invested in may improve productivity, they must be accepted and used by the employees in these organizations. This is probably the reason that adoption of new information technologies has been of significant and continuing interest to IS research, even though investigation of user acceptance of technology has been described as one of the most mature research issues in the current information systems literature (Hu, Chau et al. 1999). And this gives rise to the third research approach, which is examining the determinants of IT adoption and usage by individual users. This is of practical significance, as understanding these determinants is key to helping ensure effective deployment, leading to productivity payoffs from the investments in IT (David 1989).

In this regard, there have been several perspectives of research. One such perspective has been the implementation success at organizational level. Here, research shows that while management has been presumed to influence the extent of the adoption of an innovation by their subordinates, this influence is mitigated by certain context-specific characteristics of individual employees (Leonard-Barton and Deschamps 1988). Thus for instance, users of expert systems who had low personal innovativeness for this kind of innovation, for whom the importance of the task being computerized was low, and were low performers in their jobs perceived management as encouraging them to adopt the application. This was in contrast to those employees who rated high on all the above characteristics, who did not perceive any managerial influence in their adoption decision. The research therefore suggested that innovation diffusion was a two-step managerial process whereby employees whose traits make them more inclined to adopt an innovation do so without any managerial pressure if the innovation is availed, while those low in these characteristics typically will await some kind of pressure from management before adopting. In reality however, it may be difficult to distinguish between managerial and individual reasons.

CHAPTER 2

MODELS EXPLAINING USER ACCEPTANCE OF IS

Some of the more prominent models and theories that have been used to explain the intention to use, or usage of, information systems are summarized in the next section.

2.1 The Theory of Reasoned Action

The theory of reasoned action (TRA) is drawn from the field of social psychology. It is so named because it argues that people consider the consequences of their actions before they decide as to whether or not to engage in a given behavior. The assumption is that most actions of social relevance are under volitional control and so a person's intention to perform or not a given behavior is the immediate determinant of action. In turn, a person's intention to perform is influenced largely by two constructs: the person's attitude towards the behavior, and subjective norm. (Fishbein and Ajzen 1975).The person's attitude towards the behavior is defined as "an individual's positive or negative feelings about performing the target behavior"; while subjective norm is defined as "the person's perception that most people who are important to him think he should or should not perform the behavior in question." This theory has been used to predict a wide variety of behaviors (Sheppard, Hartwick et al. 1988; Davis 1989

2.2 The Technology Acceptance Model

The technology acceptance model (TAM) predicts information technology acceptance and usage in the workplace, and is tailored specifically to information system contexts. The model hypothesizes that two variables, namely perceived usefulness and perceived ease of use, are fundamental determinants of user acceptance of technology. Perceived usefulness is defined as "the degree to which a person believes that using a particular system would enhance his or her job performance." Perceived ease of use is defined as "the degree to which a person believes that using a particular system would be free of effort" (Davis 1989). The TAM model has been widely applied and validated with a varied set of technologies and users (Hu, Chau et al. 1999).

2.3 Motivational Model

This model identifies two types of motivation as the primary explanation for behavior (Vallerand 1997). Extrinsic motivation is defined as "the perception that users will want to perform an activity 'because it is perceived to be instrumental in achieving valued outcomes that are distinct from the activity itself, such as improved job performance, pay, or promotions'" (Davis, Bagozzi et al. 1992). Intrinsic motivation is defined as "the perception that users will want to perform an activity 'for no apparent reinforcement other than the process of performing the activity per se'" (Davis, Bagozzi et al. 1992). This model has also been used in understanding new technology adoption and use (Venkatesh and Speier 1999).

2.4 The Theory of Planned Behavior

The theory of planned behavior (TPB) is an extension of the theory of reasoned action. In addition to the attitude toward behavior and subjective norm, perceived behavioral control is the third construct that helps predict intentions to perform behaviors. Perceived behavioral control is defined as "the perceived ease or difficulty of performing the behavior" (Ajzen 1991). Moreover, the intentions to perform behaviors, together with perceived behavioral control, account for significant variance in actual behavior. In the context of IS research, this construct is replaced by "perceptions of internal and external constraints on behavior" (Taylor and Todd 1995).

2.5 Combined TAM and TPB

The combined TAM and TPB model (C-TAM-TPB) combines the perceived usefulness predictor of TAM and attitude toward behavior, subjective norm, and perceived behavioral control, all from TPB, to form the four factors that predict intentions to use (Taylor and Todd 1995).

2.6 Model of PC Utilization

Model of PC utilization (MPCU) was derived largely from the theory of human behavior (Triandis 1977) and adapted for IS contexts by (Thompson, Higgins et al. 1991), who used it to predict PC utilization. In this model, the constructs that predict usage are the following:

- Job-fit, defined as "the extent to which an individual believes that using a technology can enhance the performance of his or her job."
- Complexity, defined as "the degree to which an innovation is perceived as relatively difficult to understand and use."
- Long-term consequences, defined as "outcomes that have a pay-off in the future."
- Affect towards use, defined as "feelings of joy, elation, or pleasure; or depression, disgust, displeasure or hate associated by an individual with a particular act."
- Social factors, defined as "the individual's internalization of the reference group's subjective culture, and specific interpersonal agreements that the individual has made with others, in specific social situations."
- Facilitating conditions, which are objective factors in the environment that observers agree make an act easy to accomplish. An example in the IS context is the provision of support for users of PCs.

2.7 Innovation Diffusion Theory

Innovation diffusion theory (IDT) is grounded in sociology, and has been used extensively to study a variety of innovations. (Moore and Benbasat 1991) adopted the characteristics of innovations in (Rogers 1995) and refined a set of constructs that could be used within IS to study individual technology acceptance. The constructs found to predict technology acceptance according to this theory are:

- Relative advantage, defined as " the degree to which an innovation is perceived as being better than its precursor."
- Ease of use, defined as "the degree to which an innovation is perceived to be difficult to use."
- Image, defined as "the degree to which use of an innovation is perceived to enhance one's image or status in one's social system."
- Visibility, defined as "the degree to which one can see others using the system in the organization."

- Compatibility, defined as "the degree to which an innovation is perceived as being consistent with the existing values, needs, and past experiences of potential adopters."

- Results demonstrability, defined as "the tangibility of the results of using the innovation, including their observability and communicability."

- Voluntariness of use, defined as "the degree to which use of the innovation is perceived as being voluntary, or of free will."

2.8 Social Cognitive Theory

Social cognitive theory has the following as the main constructs that predict computer use and the use of information technology in general (Compeau and Higgins 1995):

- Outcome expectations – performance: "these are the performance-related consequences of the behavior. In particular, performance expectations deal with job-related outcomes."

- Outcome expectations – personal: "these are the personal consequences of the behavior. In particular, personal expectations refer to the individual self-esteem and sense of accomplishment."

- Self-efficacy: "this is the judgment of one's ability to use a technology like a computer for example, to accomplish a particular job or task."

- Affect: "this is an individual's liking for a particular behavior, such as use of computers."

- Anxiety: "this is the evoking of anxious or emotional reactions when it comes to performing a behavior such as use of a computer."

2.9 Unified Theory of Acceptance and the Use of Technology

The constructs used in the unified theory of acceptance and use of technology (UTAUT) to explain intention to use are synthesized from constructs drawn from all the eight different user acceptance models. The UTAUT constructs, and their composition, are described briefly.

Performance expectancy is defined as "the degree to which an individual believes that using the system will help him or her attain gains in job performance." (Venkatesh, Morris et al. 2003). The five constructs from the previous models of user acceptance that relate to performance expectancy are:

- Perceived usefulness (TAM/TAM2)
- Extrinsic motivation (MM)
- Job fit (MPCU)
- Relative advantage (IDT)
- Outcome expectations (SCT)

Effort expectancy is defined as "the degree of ease associated with the use of the system."(Venkatesh, Morris et al. 2003). The three constructs from the existing user acceptance models that depict the concept of effort expectancy are:

- Perceived ease of use (TAM/TAM2)
- Complexity (MPCU)
- Ease of use (IDT)

Social influence is defined as "the degree to which an individual perceives that important 'others' believe that he or she should use the new system." It is represented by the following constructs in the previous user acceptance models:

- Subjective norm (TRA, TAM2, TPB/DTPB, and C-TAM-TPB)
- Social factors (MPCU)
- Image (IDT)

Facilitating conditions construct is defined as "the degree to which an individual believes that an organizational and technical infrastructure exists to support use of the system."(Venkatesh, Morris et al. 2003). The three constructs included in this definition are:

- Perceived behavioral control (TPB/DTPB, C-TAM-TPB)
- Facilitating conditions (MPCU)
- Compatibility (IDT)

Performance expectancy, effort expectancy, and social influence directly influence behavioral intention to use technology, while behavioral intention and facilitating conditions both directly affect usage behavior. In addition, there are four moderating factors in this model, namely gender, age, experience, and voluntariness of use, that moderate the determinants as follows:

- Gender: performance expectancy, effort expectancy, social influence.
- Age: performance expectancy, effort expectancy, social influence, facilitating conditions.
- Experience: effort expectancy, social influence, facilitating conditions.
- Voluntariness of use: social influence.

2.10 Conclusion

While all the above models and theories have sought to explain the usage of information systems by individuals, none of them, with the exception of the TAM, has sought to explain this usage purely in terms of the intrinsic characteristics of the information system, that is, with the intrinsic qualities of the system being the primary predictor variables. And so, while they have produced very useful insight into the usage of information systems, they leave unclear the specific role or effect of the respective factors of user acceptance and, specifically, the effects that the characteristics of the information system have on usage.

Indeed, IT literature collectively suggests that user acceptance is a critical success factor for IT. It further suggests that user acceptance can be sufficiently explained, accurately predicted, and effectively managed by means of a set of relevant factors. Further, these factors are characterized by three major features: characteristics of the individual, characteristics of the technology, and characteristics of the organizational context (Hu, Chau et al. 1999).

A slightly different perspective of this is the notion that the three dimensions critical for the successful design of information systems are the software engineering dimension, the user acceptance dimension, and the incentive alignment (Ba, Stallaert et al. 2001). The software dimension mainly considers issues concerned with the features of the product itself. Along the software dimension, some of the factors that comprise the quality of software include whether:

- the software performs as intended in all circumstances.
- the system is portable to other platforms.
- the program code is well documented for future reference.
- standards of architecture and modularity have been met.
- the development effort has been cost effective.

The user acceptance dimension, on the other hand, takes the user as the focal point. Theories which exemplify this include the Technology Acceptance Model (Davis 1989), the models of cognitive fit (Vessey and Galletta 1991), and the task/technology fit.

These theories seek to explain the success or failure of the adoption of new technology from the perspective of the user.

The incentive alignment dimension is comprised of attributes that have to do with the effect of incentives and rewards on the outcome of the system. When a user's dominant use of a system and the preferred user behavior from an organization's perspective correspond, the system is described as incentive aligned. Incentive alignment is particularly pertinent in complex information systems such as distributed decision support systems, knowledge management systems, and e-business supply chain coordination systems.

CHAPTER 3

ASPECTS OF USER ACCEPTANCE OF INFORMATION SYSTEMS

The following subsections detail aspects of organizational, individual, and technology characteristics that are relevant to the study of user acceptance of information systems. These are organizational, individual, and software characteristics.

3.1 Organizational Characteristics in IS Acceptance

When IT research focuses on information systems in the context of organizations, an understanding of how organizational phenomena affect the development and use of technologies and, conversely, how technologies impact on organizations, is of major interest (Orlikowski and Barley 2001).

As part of the organizational aspects that influence user acceptance of technology is the social influence of external inputs, such as communicated information, on individuals' attitude. As influential as TAM has become for explaining acceptance of IS, the omission of subjective norm from it was identified as an important area in need of further investigation (Davis 1986; Davis, Bagozzi et al. 1989). Changes in attitudes and actions as a result of this social influence may occur at different levels (Kelman 1958). These levels correspond to differences in the processes by which the individual accepts the influence. There are three distinct processes of social influence that shape individual behavior, namely compliance, identification, and internalization. These are described as follows:

Compliance refers to the situation whereby the individual adopts the induced behavior with the expectation of gaining rewards or avoiding punishment as the primary motive, rather than because of a belief in the contents of the behavior. Identification on the other hand refers to a situation whereby an individual accepts influence mainly as a means of satisfying a self-defining relationship with another individual or group. Internalization occurs when an individual accepts influence because it agrees with his value system (Kelman 1958).

By distinguishing between these three processes, it is possible to determine if usage behavior is caused by external influence or by one's own attitude. Each of processes has a distinct set of causal conditions corresponding to a characteristic pattern of internal responses involving thoughts and feelings in which the individual engages while adopting the induced behavior. Similarly, each of the three processes gives rise to a distinctive set of consequent conditions that have a bearing on the induced response. Thus for example, behavior induced through compliance tends to be performed under the surveillance of the influencing force, whereas behavior induced through identification tends to be performed in the context of one's relationship with the agent. When the behavior is brought about through internalization however, it tends to be performed under conditions that the individual deems relevant, regardless of the surveillance or salience (Kelman 1958).

In the context of usage of a new information system, these social influence processes determine the individual user's commitment to the use of new information technology (O'Reilly and Chatman 1986). In contrast to the common idea of use in terms of a dichotomy of use versus non-use, the concept of social processes suggests that the use of an information system is a continuum that spans from avoidance of use (or nonuse), through unenthusiastic use (compliant use), to enthusiastic and consistent use (from internalization) (Malhotra and Galletta 1999).

A different perspective of research shows that while management has been presumed to influence the extent of the adoption of an innovation by their subordinates, this influence is moderated by certain context-specific characteristics of individual employees (Leonard-Barton and Deschamps 1988). Thus for instance, users of expert systems who had low personal innovativeness for this kind of innovation, for whom the importance of the task being computerized was low, and were low performers in their jobs, perceived management as encouraging them to adopt the application. This was in contrast to those employees who rated high on all the above characteristics, who did not perceive any managerial influence in their adoption decision. The research therefore suggested that innovation diffusion was a two-step managerial process whereby employees whose nature makes them more inclined to adopt an innovation do so without any managerial pressure if the innovation is availed, while those low in these characteristics typically will await some kind of pressure from management before adopting. Here, it is difficult to distinguish between managerial and individual reasons. End-user evaluations of technological innovations are affected by the attributes, opinions, and actions of the leaders of their organizations (Chakrabarti 1974; Leonard-Barton 1988). They also heavily depend on individual users' personal interests, needs and skills. It is therefore reasonable to assume that implementing an innovation of high complexity, which is one that many members of the organization must use for it to be of benefit to that organization (Leonard-Barton and Deschamps 1988), involves internal diffusion by the users even after the organizational authority has okayed the decision to adopt. It is also reasonable to assume that in the diffusion of innovations within an organization, subsequent to the adoption decision by management, there exists the possibility of management behavior impacting on the individual decisions by the users to adopt or to continue using the innovation by means of a variety of forms of influence varying from express directives to covert signals of support (Leonard-Barton and Deschamps 1988).

One of the theories that include organizational perspective in the study of information technology is the adaptive structuration theory (AST). It explains the interaction of groups and organizations with information technology (Desanctis and Poole 1994). As groups and organizations use information technology for their work, they continuously have perceptions concerning the role and utility of the technology, and how it applies to their work. These ideas vary widely with groups, and affect the manner in which the technology is used, and therefore how much it affects group outcomes. Some of the notable points of AST are that:

- It deals with the evolution and development of groups and organizations.
- It views groups or organizations as systems that have observable patterns of relationships and communicative interaction among people, creating structures.
- Systems are produced by actions of people creating structures, which are sets of rules and resources.
- Systems and structures exist in a dual relationship with each other in the sense that they tend to produce and reproduce each other in an ongoing loop. This is what is referred to as the "structuration process."
- This structuration process can be very steady, or it can change substantially over time.

3.2 Individual Characteristics in IS Acceptance

Decision making is commonly acknowledged as one of the most important human skills. Perhaps as a consequence, it is the subject of many different theoretical frameworks. In the context of user acceptance, decision-making is important because it is the mechanism by which users evaluate or assess software so as to arrive at a judgment regarding its acceptability. There are two major research areas of interest in this regard. One is decision-making as a mechanism of conflict resolution. This is mainly espoused by social psychologists (Janis and Mann 1977). The second research area of interest focuses on the frequently observed departure from rational choices in forming judgment. Such judgment is usually based on heuristics, which are informal methods or experience, and frequently employs a form of trial and error (Kahneman, Slovic et al. 1982). These two areas are further discussed below.

3.2.1 Decision Making for Conflict Resolution

It is, in general, widely accepted that decision-making consists of a number of steps such as recognition, formulation, and generation of alternatives, information search, selection, and action. Further, every decision is held to have three components:-

- Criteria: The standards by which decision makers evaluate the alternatives.
- Alternatives: Specific courses of action or options under consideration.
- Cause and effect beliefs: These are cognitions linking specific assumptions, beliefs or alternative attributes. These cognitions are also referred to as models, assumptions, beliefs, or theories.

In addition, a decision-making model construct is a model that determines the process by which a decision is made. It has three major elements, namely:

- How criteria are determined- that is which criteria should be used in making a decision. This includes dealing with multiple, conflicting criteria.

- How alternatives are generated.

- How the alternatives are evaluated against criteria. This involves how cause/effect relationships are established, and how cause/effect conflicts are resolved.

As a result of an extensive review of the literature on effective decision-making, (Janis and Mann 1977) suggest seven major criteria for determining whether decision-making procedures are sound. Decisions meeting these seven criteria have a higher chance of meeting the decision maker's objective than those that do not. The seven criteria are that the decision maker:

- Thoroughly canvasses a wide range of alternative courses of action.

- Examines the full range of objectives to be fulfilled and the values implicated by the choice.

- Carefully takes into account the costs and risks of negative consequences as well as the positive consequences that could be attendant in each alternative.

- Aggressively searches for new information that is relevant for evaluating the alternatives further. Correctly interprets and objectively considers any new and relevant information regarding the task of selecting from among the alternatives.

- Reexamines all the consequences, positive and negative, of all alternatives before making a final decision.

- Makes detailed arrangements for implementing the chosen alternative with contingency plans for any risks that might materialize.

The effectiveness of the decision-making construct may be gauged by whether the construct allows or encourages:

- The decision-maker to include all relevant criteria.

- The consideration of all alternatives.

- The decision maker to access the most accurate information regarding cause and effect.

Three strategies identified in the literature for decision making are:

- Maximizing-minimizing
- Satisficing
- Singling out a criterion for maximization

Under maximizing-minimizing, the decision-maker develops a utility index function based on certain criteria, which he then maximizes or minimizes. Each alternative is then scored on each criterion, and a weight determined for each criteria. An index is then computed by an addition of the product of the criteria score and weight for each alternative, and picking that with the maximum or minimum as desired. This strategy is also referred to as optimizing (Janis and Mann 1977). However, people rarely adopt this approach to decision making. One reason is the vast amount of information that the decision maker might have process as he attempts to obtain the level of knowledge necessary, and which might lead to information overload. Also, more variables might have to be taken into account than can be kept in the mind at the same time. The number of these relevant pieces of information might exceed seven to nine, which is the normal limit of man's capacity for processing information in the immediate memory. Other factors have been identified that limit the use of this seemingly ideal optimization strategy. These include the inability of individuals to handle the complicated mathematics that would be necessary for optimization; the existence of other nonquantifiable variables that nevertheless bring about a deviation from maximization; the force of habit; the effect of tradition; and the influence of social institutions. The consequence of all these may be a decision that may not appear "rational" (Janis and Mann 1977).

Satisficing: Under this strategy, the decision maker sets the minimum acceptable level for each criterion. Each alternative is then evaluated against the set minima, and those alternatives not meeting the criteria are eliminated. The strategy is alternatively called the 'good enough' strategy (Simon 1997), since the alternative only needs to be good enough or acceptable to be adopted. This strategy appeals to the fact that many people prefer the simplification of complex decision problems. In the event that no alternatives meet the minimum criteria, the following options are followed-

- Criteria are eliminated in increasing order of importance until an acceptable alternative is obtained.
- The minimum acceptable levels are lowered until an acceptable minimum alternative is found.
- The search for a suitable alternative is carried out until one is found that fits the set of criteria.

In the event that there are several alternatives that meet the criteria set, the following options can be pursued:

- Choose a particular criterion or subset of criteria to maximize or minimize.
- Include additional criteria that will differentiate between the alternatives.

Singling-out a criterion for maximization: Here, the decision-maker selects what is considered to be the most important criterion and uses this to select the most suitable alternative as the one maximizing this criterion. It could actually be taken as the simplest form of satisficing and is also referred to as quasi satisficing (Janis and Mann 1977).

The satisficing strategies can be contrasted to optimization on at least the following four dimensions (Janis and Mann 1977):

- Number of requirements that must be met: In determining whether or not to change to a new alternative, the decision maker depends on a small number of requirements, and in some cases only one. He deliberately disregards other values that could be relevant to the decision. In contrast, while using the optimizing strategy, he considers a much larger set of requirements or objectives, and determines the alternative that scores highest with these requirements.

- Number of alternatives generated: Employing the satisficing decision rule, the decision maker sequentially examines each alternative that he has, and settles on the first one that proves satisfactory. Thus, relatively few alternatives are looked for or considered. In contrast, with the optimizing strategy as many alternatives as possible are considered.

- Ordering and retesting of alternatives: With the satisficing strategy, the decision maker typically examines his alternatives as they come, rather than in any particular order, until he gets one that meets the minimum requirements. In contrast, with the optimization strategy, the alternatives are ordered in a way that facilitates comparative judgment to enable picking of the best.

- Type of testing model used: using the satisficing decision strategy, the decision maker will usually only check to see whether an alternative meets the minimum that has been set for a requirement. If there are multiple requirements, their minima are all given the same weighting, and the alternative is checked to see if it meets these. In contrast, the model used for optimizing is usually a weighted additive model. Thus, in carrying out an evaluation of alternatives, account must be taken of all the levels of the benefits and costs of each objective or factor as well as the relative importance of each. This enables the consideration of tradeoffs whereby high levels of a factor that is deemed important are obtained while permitting relatively low levels of factors considered less important.

Furthermore, criteria used in decision making can be grouped into value-based or need-based criteria. Value-based criteria refer to the preferences of the decision maker. Since the weights are subjective decision maker values, there is no empirical technique for determining weights. Thus, these criteria are better suited for satisficing strategies. Predictive criteria on the other hand are standards used in decision making to predict a desired preference. And since maximizing strategies work better with empirically determined criteria weights, predictive criteria are better suited for these kinds of strategies.

3.2.2 Decision Making using Heuristics

In certain situations, rather than utilize rational reasoning in evaluation, people have been shown to resort to intuitive judgment. In so doing, they employ a limited number of heuristics that reduce the complex task of assessing probabilities and values to simpler forms of judgment. These, while they have some validity and can therefore be useful, do sometimes lead to serious and systematic errors which are predictable, however. Some of the common heuristics and the biases they lead to are summarized below.

The representativeness heuristic: This heuristic refers to instances where the probability of the occurrence of an event A is evaluated by an individual on the basis of the similarity of A to another event or process B, rather than by available empirical statistics that are actually available about the probability of the event. Thus, when A is judged to be representative of B, the probability that it originates from B is judged to be high, and conversely, when it is judged not to be representative of B, it is judged to be low. The fallacy from using this heuristic arises from the fact that similarity or representativeness is not necessarily influenced by the same factors that should go into the judgment of probability.

The availability heuristic: This heuristic refers to situations where people assess the frequency of a class or the probability of an event based on the ease with which instances or occurrences can be recalled. It is a useful rule for estimating frequency or probability to the extent that instances of large categories of objects are recalled better and easier than those of classes that occur less frequently. However, ease of recall may be due to certain other factors that are not related to the probability of occurrence. Such factors include familiarity and the salience or vividness with which a similar event is recalled. Also, recent occurrences are likely more easily recalled than those that had occurred earlier. Cases where such factors are the main determinant in estimating the probability of an event result in bias, since these factors in reality have no bearing on the likelihood of an event.

The anchoring and adjustment heuristic: This heuristic refers to those situations whereby people make estimates by having an initial value, which is referred to as the anchor, and then making adjustments in order to arrive at the final estimate. The anchor is an implicitly suggested reference point. Adjustments are typically insufficient. Moreover, the final estimates are usually biased towards the initial anchor, so that different anchors lead to different final estimates which are usually closer to the respective anchor than to any other, hence the bias.

3.3 Technology Characteristics in IS Acceptance

Technology characteristics have been investigated in a number of models. These are briefly discussed below.

3.3.1 TAM Model

The goal of the Technology Acceptance Model (TAM) is to provide a general explanation of the determinants of computer acceptance that is capable of explaining the behavior of users over a wide range of end-user computing technologies and user populations, while having theoretical backing, and remaining parsimonious at the same time (Hu, Chau et al. 1999). The TAM is based on the theory of reasoned action (TRA) (Fishbein and Ajzen 1975). According to TRA, beliefs influence attitudes, which in turn lead to intentions which then guide or give rise to behaviors. TAM adapts this relationship into an IT user context. In this context, both perceived ease of use and perceived usefulness influence the attitude towards an application. This attitude in turn influences the intention to use the application. In addition, perceived ease of use has an effect on the perceived usefulness, while perceived usefulness directly affects the intention to use the technology. The model is depicted diagrammatically below:

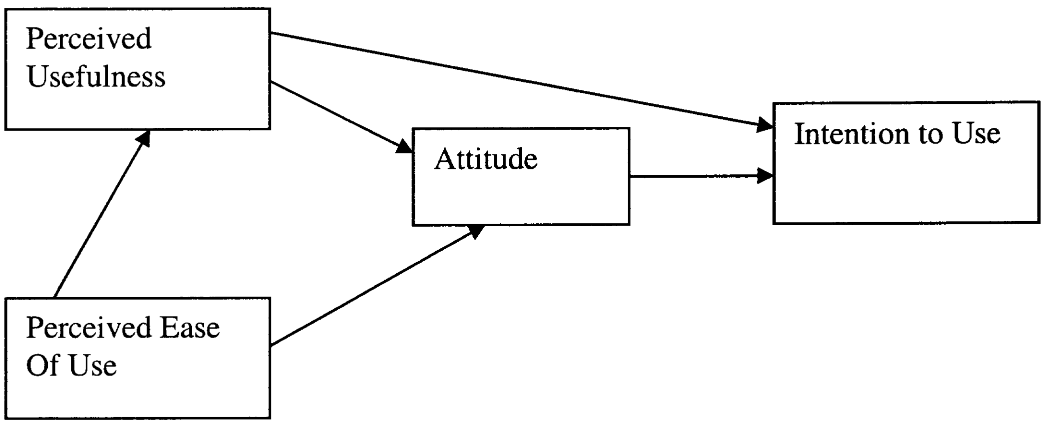

Figure 3.1 The TAM model.
(Source: Hu, Chau et al. 1999)

TAM has been investigated in a wide variety of user group, technology, and organizational settings. While perceived usefulness has consistently been identified as important in the formation of attitude, the results for perceived ease of use have been less consistent and significant. A plausible explanation in the literature is that importance of perceived ease of use as a determinant of the intention to use a technology diminishes with the users' prolonged exposure to the technology (Hu, Chau et al. 1999). Nevertheless, TAM is has accumulated considerable satisfactory empirical support for its overall explanatory power, and has suggested individual causal links across a wide variety of technologies, users and organizational contexts.

In comparison with the other models and frameworks, the advantages of TAM include its being parsimonious, having a strong theoretical basis, its significant empirical support, and most important, its being IT specific. It has, therefore, become a dominant model for investigating technology acceptance by users. However, (Hu, Chau et al. 1999) found that the model had not been sufficiently tested with professionals such as physicians in their particular professional contexts. Of research interest was the possible difference between such a subject group and others such as students in such factors as general competence, intellectual and cognitive capacity, specialized training, and professional work. They used the TAM model to investigate the acceptance of telemedicine. Telemedicine is an IT based innovation that aims to support and improve the care that physicians provide to patients as well as improve the competitiveness of health care organizations. They found that while perceived usefulness had a significant and strong influence on the physicians' intention to adopt the technology, perceived ease of use did not. Perceived ease of use and perceived usefulness explained only 37% of the variances in attitude towards the technology, while perceived usefulness and attitude together only explained 44% of the variances in the intention to use the technology. The results suggested that other factors should be included to the research model. They concluded that the fact that the TAM appeared to lack enough explanation for the attitude and intention of physicians might have to do with the characteristics of healthcare or the very nature of the profession (Hu, Chau et al. 1999). This finding puts in question the completeness of the TAM model for explaining user acceptance of technology in terms of the technology's intrinsic qualities. Perhaps there are other qualities of technology that are of particular importance to people in certain professions. Based on general literature in the field of healthcare, security and privacy could be two such factors that could be particularly important in that industry.

Therefore, a major weakness of the TAM model is that there is no theory that shows it to be exhaustive in explaining user acceptance in terms of all the relevant qualities of the information systems.

3.3.2 Revised ISO Model

A review of the existing literature reveals that there are other software qualities that have a bearing on the user acceptance of applications. Indeed, the software marketplace has become much more competitive, and the software systems correspondingly more complex. Moreover, customer needs are constantly evolving, and fewer applications are operating in isolation today. As a consequence, software quality is now rarely evaluated in terms of functionality alone. Rather, non-functional qualities such as portability, modifiability, and extendibility have become increasingly important for successful applications (De Simone and Kazman 1995). The International Standards Organization standard ISO 9126-1 defines quality as 'a set of features and characteristics of a product or service the bear on its ability to satisfy stated or implied needs.' The ISO 9126-1 quality model is now widely accepted as a modern product quality specification. It suggests six high-level quality characteristics. This are reproduced below from (Losavio, Chirinos et al. 2004).

- Functionality: The capability of the software product to provide functions that meet stated and implied needs when the software is used under specified conditions.

- Reliability: The capability of the software product to maintain its level of performance under stated conditions for a stated period of time.

- Usability: The capability of the software product to be understood, learned, used, and attractive to the user, when used under specified conditions, or the effort needed for use.

- Efficiency: The capability of the software product to be modified.

- Maintainability: The capability of the software product to be modified, where modifications could include corrections, improvements, or adaptations of software to changes in the environment, requirements, or the functional specifications.

- Portability: The capability of the software product to be transferred from one environment to another, where the environment may include organizational, hardware, or software environments.

Rivard, Poirier et al. 1997 further refined the ISO model, resulting in a model with eight dimensions. This model, together with the criteria that are proposed to measure each dimension, is reproduced below in Figure 3.2. Along with each dimension, it shows the key aspects that form the criteria that for that dimension.

Thus this model is actually an augmentation of the ISO model. The following dimensions given below, together with their definitions, are added:

- Economy: This dimension is relevant in the context of end-user computing, and a critical aspect is the profitability to a user of using the application.

- User-friendliness: the ease of learning how to use a system, how to operate it, and how to prepare the input data, how to interpret the results, and how to recover from errors.

- Understandability: This is the extent to which one can understand what an application does, including its structure and its modules.

- Verifiability: This is the ease of testing the application to ensure that it performs its intended function.

It should however be noted that ISO 9241-11:1998 (Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability), defines usability as "the extent to which a product can be used to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use." This definition of usability includes effectiveness, which is an aspect of functionality. This demonstrates the inconsistency in the definition of performance criteria in the literature even within the same standards organization.

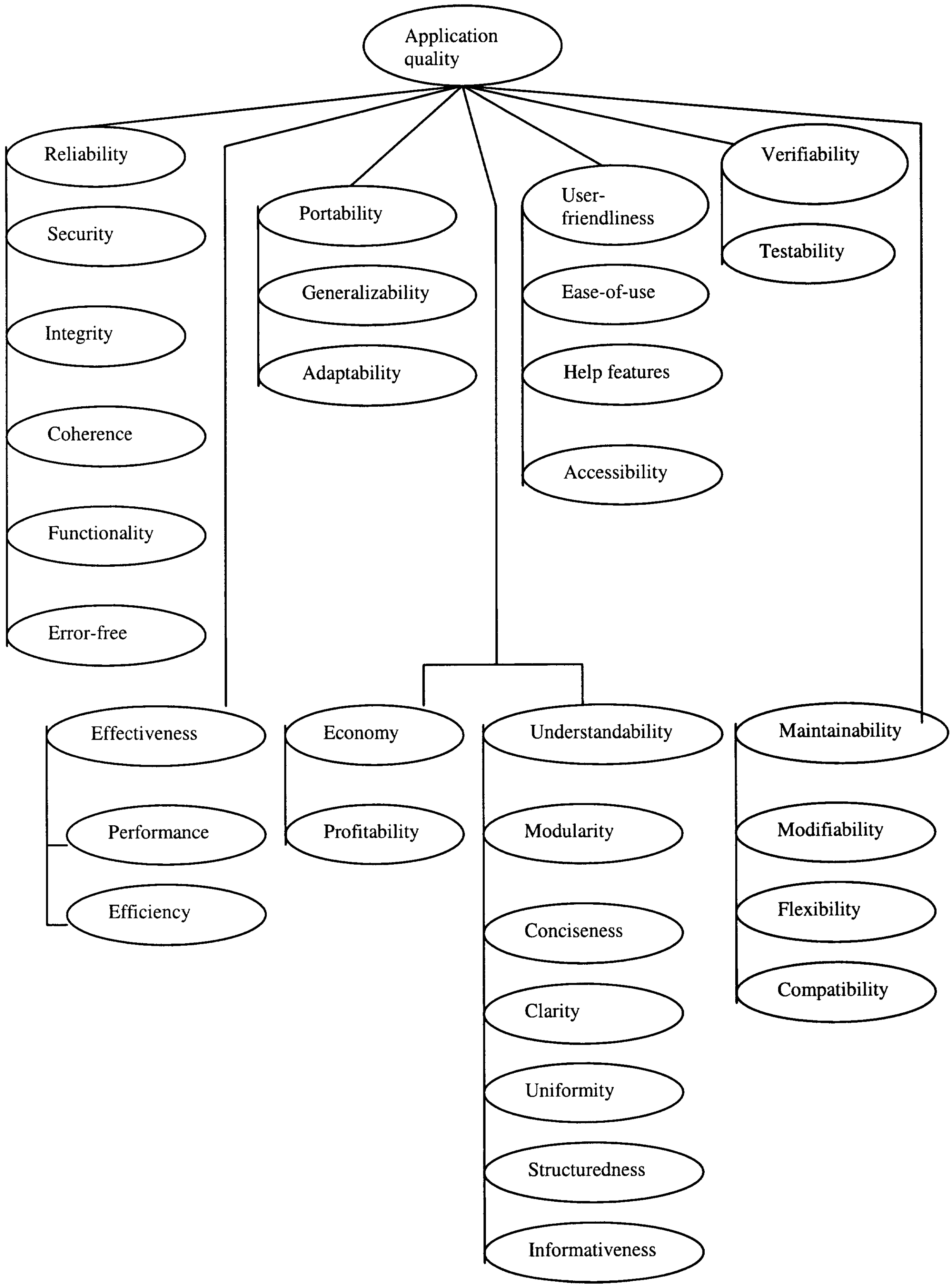

Figure 3.2 Revised ISO model.
(Source: Rivard, Poirier et al. 1997)

These two models are very useful in providing further insight as to the intrinsic qualities of applications that are of importance to user acceptance. They clearly show that there are other software attributes, apart from usability and usefulness, that impact on user acceptance of an application. However, a major weakness is that they do not derive from any particular theory. Rather, they rely on a survey of existing literature to come up with the concepts of software qualities. It is therefore difficult to discern whether they completely describe all the factors relevant for user acceptance of software, or at any rate, the extent to which they do so. A key problem in doing such a survey of the literature is that different attributes of quality seem to share the same meaning, while a given term could have different meanings, or be used differently (Rivard, Poirier et al. 1997).

3.3.3 Software Engineering Model

In the field of software engineering, software qualities are often classified into functional requirements, non-functional requirements, and domain requirements (Sommerville 2000).

Functional requirements refer to the services that the system should provide, and may also state what the system should not do. Non-functional requirements refer to the constraints imposed on the services or functions of the system. They are not directly concerned with the specific functions delivered by the system, and may relate to system properties considered to be emergent, such as reliability. Domain requirements are the requirements specific to the domain to which the application belongs, and reflect the character of that domain. They may be functional or non-functional. Non-functional requirements can be further classified into product requirements, which specify product behavior; organizational requirements, which are derived from organizational policies and procedures; and external requirements, which include requirements that are derived from factors external to the system and its development process such as legislative and ethical requirements. A diagrammatic representation of the nonfunctional requirements is reproduced in Figure 3.3 below from (Sommerville 2000).

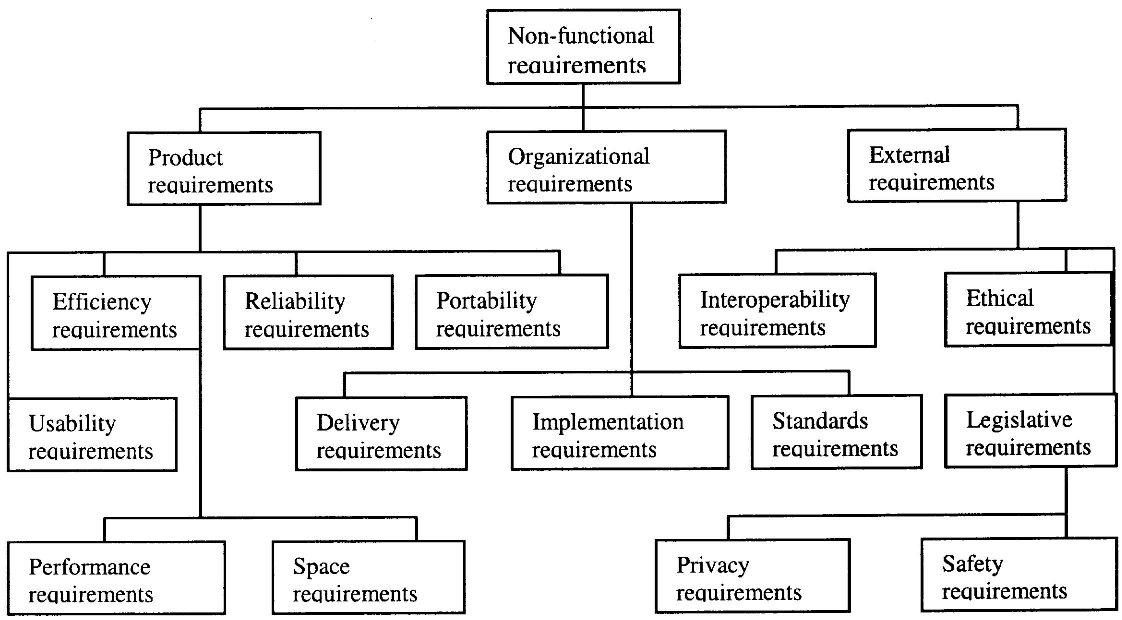

Figure 3.3 Non-functional requirements.
(Source: Sommerville 2000)

However, the distinction between the three broad classes of software requirements is not so clear-cut in reality (Sommerville 2000). Moreover, even though the classification is very useful in helping understand software quality, it does not derive from any theory, and so it is difficult to comment on its exhaustiveness and parsimony.

3.3.4 The Dependability and Security Model

The classical model of dependability is as a composite characteristic comprised of the attributes of reliability, availability, safety, and security (Laprie 1992). Security on the other hand is described by three aspects, namely confidentiality, integrity, and availability (EC 1993).

Elsewhere in the literature, computing systems may be viewed as characterized by four basic concepts, namely functionality, performance, cost, and dependability (Avizienis, Laprie et al. 2004). In this environment, dependability is the ability to deliver trustworthy service, while service refers to the system's behavior as perceived by its user or users. In this respect, a user is another system that interacts with the computer system in question at the interface of the service. The function of a system is the intended purpose of that system as described by the system specification.

Dependability is taken to be a compound concept that includes availability, reliability, safety, integrity, and maintainability. In this context, availability is the readiness for correct service; reliability is the continuity of correct service; safety is the absence of serious consequences to the users and the environment; integrity is the absence of improper alterations of system states; while maintainability is the ability to undergo repairs and modifications (Avizienis, Laprie et al. 2004).

The related concept of security has been described as the concurrent existence of availability (only for authorized users), confidentiality, and integrity (where 'improper' means 'unauthorized' in the previous definition (Avizienis, Laprie et al. 2004), and confidentiality, which is defined as the absence of unauthorized disclosure of information.

There exist various versions of the given definition of security. In some cases, extra aspects such as denial-of-service and authenticity are included, while in other cases, a different grouping of aspects is given, as in (ISO 1989; Muftic 1989).

There also exists a formal model of confidentiality as a description of information flow in a secure system that is aimed at identifying paths that could lead to inappropriate disclosure of information (Bell and LaPadula 1973).

Here it can be seen that dependability is seen as distinct from, and encompassing, reliability, which are viewed as synonymous in other parts of the IS literature. Moreover, dependability and security share the attributes of availability and integrity. This is a slight modification of the earlier situation where security was given as an attribute of dependability. Also, security is in some cases taken to include confidentiality. All this is a further illustration of the confusion in the literature concerning the concepts defining system properties.

3.3.5 Mainframe Computer Performance Factors

The fact that there are other software performance factors apart from just usability and usefulness is borne out by a consideration of the factors that are generally regarded as responsible for the popularity of the mainframe computers. The mainframe continues to be the foundation of modern business in areas as diverse as banking, finance, healthcare, insurance, utilities and government among others. All these businesses rely on the mainframe to:

- Perform large-scale transaction processing, running into thousands of transactions per second.

- Support myriad users and application programs concurrently accessing many resources.

- Manage stupendous amounts of information in databases.

- Handle large bandwidth communications.

Mainframes are able to accomplish all this because of their inherent reliability, stability, security and scalability. These qualities also lead to the popularity of mainframes in many IT organizations for hosting the most important mission-critical applications. Such applications include customer order processing, financial transactions, production and inventory control, and payroll (IBM).

Reliability: In particular, with regards to mainframe computers, reliability, availability, and serviceability are grouped together and are seen as a very important set of qualities in data processing (IBM). These terms are described as follows:

- Reliability: This involves the use of high-quality hardware and software components, and extensive self checking and self-recovery by hardware components.

- Availability: This is the ability of hardware to detect and automatically replace failing hardware elements, and for system software to detect, isolate, and recover failing software components.

- Serviceability: This involves the non-disruptive servicing or upgrading of hardware and software with minimal impact to the operational system through well-defined units.

Security: In an IT environment such as mainframe computing, data security is defined as protection against unauthorized access, transfer, modification, or destruction, whether accidental or intentional. Thus, a secure computer system prevents users from accessing or changing any objects on the system, including user data, except through system-provided interfaces that enforce authority rules (IBM).

Scalability: In the information technology industry, change is constantly occurring. For instance, in business, good results can set off a growth in the IT infrastructure to cope with the resulting increased demand. The degree to which the IT organization can add capacity without suffering disruption to its normal business processes or without suffering excessive overhead in form of nonproductive processing is to a large extent determined by the scalability of the particular IT infrastructure.

In this respect, scalability can be defined as the ability of the hardware, software, or a distributed system to continue to function well even as it is changed in size or volume. An example of this is the ability to retain performance levels when adding processors, memory and storage. Similarly, a scalable system efficiently adapts to work, with larger or smaller networks performing tasks of varying complexity.

3.3.6 E-Commerce Software Performance Factors

A set of quality dimensions has been proposed for e-commerce software tools (Krishna and Subramanyam 2004). These are capability, usability, performance, reliability, and documentation. The respective definitions of each of these dimensions in this context are given below.

Capability: This refers to the product functionality in terms of key features supported by the software product. Included in this is the customizability of the software to the customer's existing infrastructure, which is expected to be high.

Usability: This is the effort spent by users of the software in learning how to use the software. It includes a consistent and intuitive user interface, as well as such features as on-line help and context-sensitive help. The greater the number of features provided by the software however, the greater the effort usually necessary by users to learn the features.

Performance: This attribute indicates such aspects as response time to users' operations, time to display outputs, and efficient use of computing resources such as memory and storage space for user operations.

Reliability: This software attribute represents the conformance quality. For the user, it is a measure of the amount of disruption during usage, which should be kept to a minimum. For e-commerce software, a very high level of reliability, with minimum downtime - to the extent that users take the service for granted - is required.

Documentation: The key aspects of this attribute are clarity and ease of access. The quality of documentation and the accompanying manuals for software has an impact on the user's ease and speed of learning.

3.3.7 Crosscutting Quality Attributes Approach

Non-functional requirements can be described as quality attributes, such as response time, accuracy, security, and reliability which affect the system as a whole (Moreira, Araujo et al. 2002). A major problem with dealing with these quality attributes is that they are very diverse. Also, the fact that in most current approaches of software engineering these attributes are treated separately from functional requirements points to the difficulty in achieving the integration of the system's requirements. Another disadvantage of these approaches is their failure to address the fact that some of these attributes can simultaneously affect several requirements.

A proposed approach accomplishes the integration of requirements by 'weaving' the quality attributes with the functional requirements using the concepts of overlapping, overriding, and wrapping.

In overlapping, the quality attribute requirements modify the functional requirements that they crosscut. The quality attribute requirements may be required before or after the functional requirements. With overriding, the quality attribute requirements override the functional ones they crosscut. Thus, the system behavior described by the non-functional requirement substitutes that of the functional requirement. In the case of 'wrapping', the quality attribute requirements "encapsulate" the functional requirements they crosscut. This means that the functional requirements must be fulfilled 'within' the specifications of the quality attribute. A major drawback of this approach is that there is as yet no consensus as to what the crosscutting quality attributes are (Moreira, Araujo et al. 2002).

3.3.8 WOSP Model

Whitworth and Zaic, 2003, suggest yet another model for explaining those qualities of software systems that are relevant for user acceptance. This model, called the Web of System Performance (WOSP) model, is derived from systems theory. Systems theory has been defined as “the transdisciplinary study of the organization of phenomena, independent of their substance, type, or spatial or temporal scale of existence” (Principia Cybernetica online encyclopedia). It investigates both the principles common to all complex entities, and the models which can be used to describe them. And in common with all other complex systems found in nature, the performance of information systems involves multiple aspects. A system’s performance in this context is defined as how successfully it interacts with its environment. The aspects can be grouped into the basic system elements of boundary, internal structure, effectors, and receptors. Each of these elements in turn has a dual role in system performance that seeks to maximize opportunity or value, and minimize risk or loss. This gives rise to eight system performance goals, which are the fundamental attributes of the system. The elements and their system performance purposes or goals are:

- Effectors:
  - Functionality – to act on the environment
  - Usability – to reduce the cost of any action
- Boundary:
  - Security – to prevent unwelcome entry
  - Extendibility – to use outside objects or materials
- **Structure:**
  - Reliability – to continue performing the same despite internal change
  - Flexibility – to perform differently given external change
- Receptors:
  - Connectivity – to exchange social meaning
  - Confidentiality – to control or limit social meaning exchange

Computer systems are regarded as systems in the general sense at several levels (Churchman 1979). At the lowest level with the computer being considered as hardware, a computer can be viewed as a mechanical system. In terms of the above system elements, the casing would be the boundary, the keyboard and mouse would be the receptors, and the printer and screen would be the effectors, while its internal architecture would give the internal structure.

The next higher level comprises the hardware together with a software system. Here, the system boundary separates system memory from non-system memory for example, across which input and output data access and depart the system. The software has internal structure, whereby parts call as well as comprise each other. Subroutines processing input would be examples of receptors, while driver programs would be examples of effectors acting on the system's environment (Whitworth and Zaic 2003).

The next higher level of system is the human-computer combination, which now includes human cognitive processing. And at the highest system level, a computer-mediated community can be considered a socio-technical system, with social processes that can be modeled.

The WOSP model further predicts that the extent to which each of these performance goals is important depends on both the system and its environment. In the context of information systems, these goals can be viewed as system performance factors or criteria.

The WOSP model can be represented diagrammatically as shown below in Figure 3.4, reproduced from (Whitworth and Zaic 2003).

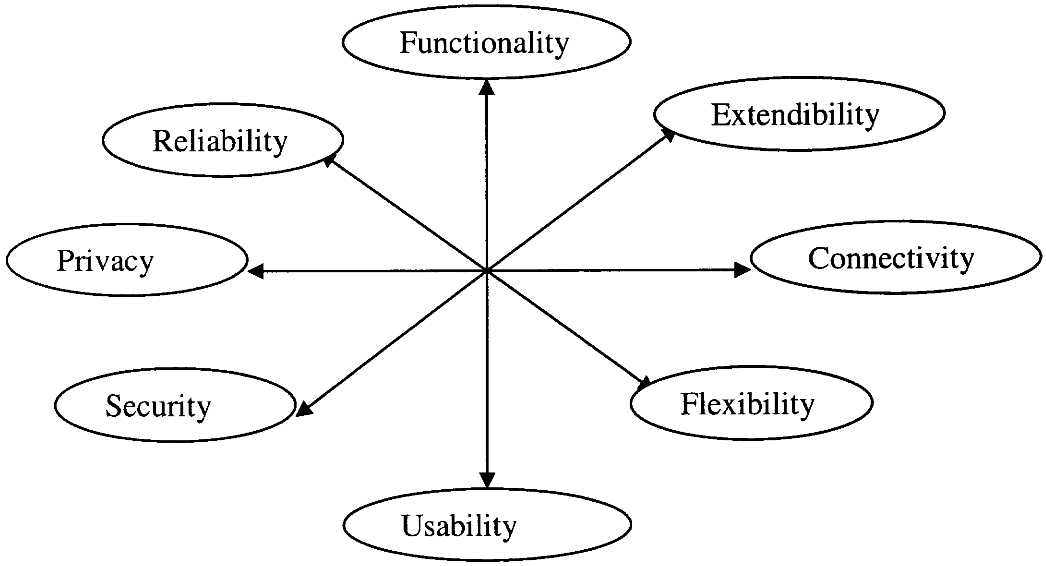

Figure 3.4 The WOSP model.
(Source: Whitworth and Zaic 2003)

Thus, a major difference between the WOSP model and the others that have been proposed in the literature to explain the qualities of software that have a bearing on user acceptance is that it is based on a well-known theory, and is exhaustive in its description of the software qualities. A check on its generalizability can be made by superimposing it on the other models earlier presented. This is done below in Figures 3.5 and 3.6. The WOSP quality that corresponds to an intrinsic quality of a given model is shown in brackets.

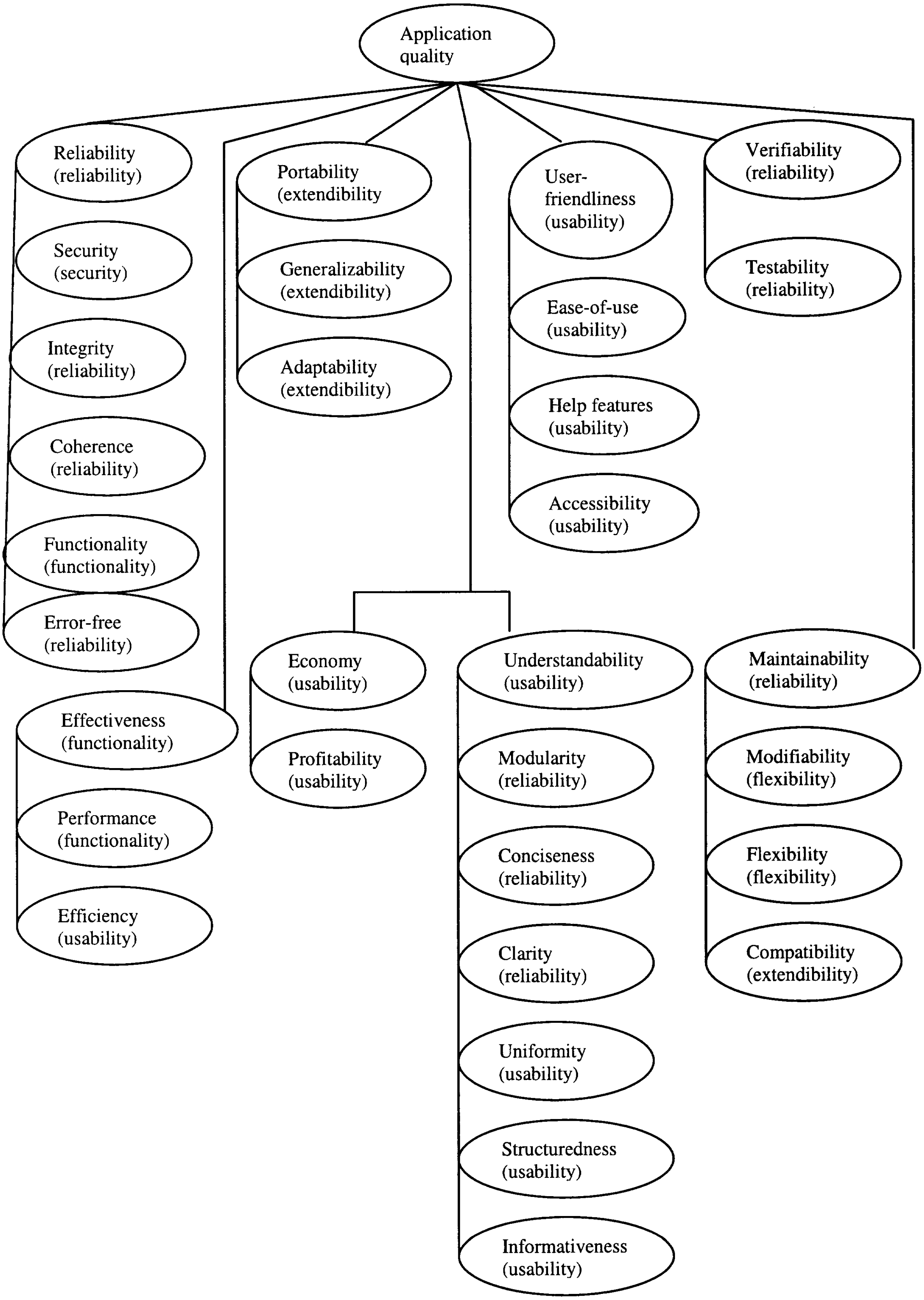

Figure 3.5 Superimposition of WOSP model on the revised ISO model.

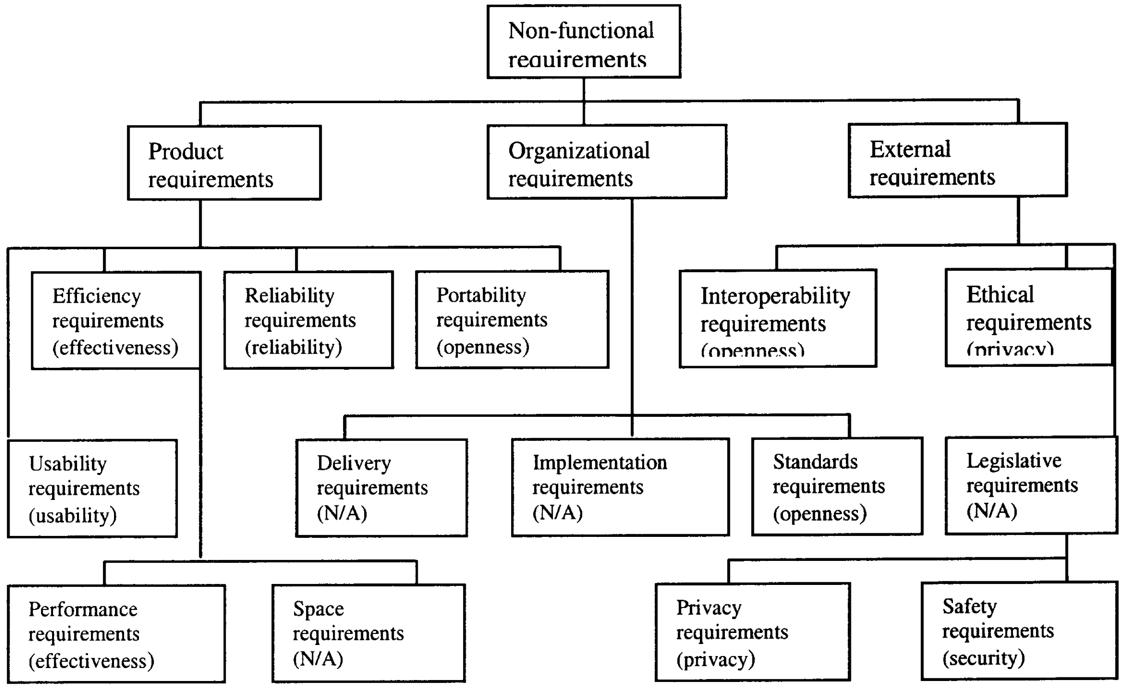

Figure 3.6 Superimpostion of the WOSP model on the software engineering model.

The fact that the WOSP model covers the other models means that it is at least as comprehensive, and is also parsimonious in its attributes. This makes it a good candidate for further research as to its validity. It is noteworthy that currently, the challenge of evaluating software with respect to their non-functional qualities is compounded by the fact that there is no unit of measurement of these qualities, or an objective manner of evaluating them. Also, there are differences in how the non-functional qualities are interpreted by different individuals. Moreover, non-functional features have not been completely defined authoritatively (De Simone and Kazman 1995).

3.4 Conclusions

Organizational characteristics undoubtedly exert an influence on how users relate to software. Organizational issues are however largely external, or pertain more to the environmental context within which the user interacts with the software. Moreover, the organizational characteristics do not vary much within a given organizational entity. They are therefore relatively easy to manipulate and adjust for in real life, or to control for in an experimental situation.

In contrast, individual characteristics are intrinsic to people. It is difficult to control for them, especially given that they likely vary from individual to individual. Account would, therefore, have to be taken of, and allowances made for, these individual differences both in real life and in an experimental situation.

Regarding technological characteristics, there obviously are several factors intrinsic to software that influence the user's perception of the performance of the software. It is also apparent at this point that there is a lot of conceptual confusion in the literature with respect to the software performance constructs and accompanying factors, with the constructs and factors being defined differently in different models. It is therefore essential at this juncture to review the literature in detail so as to gain a better understanding of the constructs and factors involved in software performance, and with a view to resolving the existing inconsistencies.

The WOSP model is the most comprehensive in terms of inclusion of software quality factors. It also offers an explanation of how these factors relate to each other. This model therefore offers the best framework for further investigating the technological characteristics of software.

CHAPTER 4

LITERATURE REVIEW OF SOFTWARE PERFORMANCE FACTORS

This review gives an overview of the existing information in the IS literature concerning the eight software performance factors contained in the WOSP model, namely flexibility, reliability, security, extendibility, privacy, connectivity, functionality, and usability.

4.1 Flexibility

4.1.1 Significance

Today, there is increased attention being paid to the variation of work activities and the concomitant need for flexibility. This is being reflected by a growing concern for tailorability issues, and contrasts with the earlier emphasis of conventional approaches on general aspects of work activities (Kjaer and Madsen 1995). Two major reasons cited for the need for tailoring are diversity and fluidity. Tailorability involves the modification and adaptation of systems after they have been put into use.

A recent survey found that the creation of a strong and flexible IT infrastructure was the top priority among 150 managers of IT who were polled (InformationWeek 1999). IT infrastructure is generally understood to be made up of two components (Byrd and Turner 2000): the human IT infrastructure, which comprises the experience, competencies, commitments, values, and norms of the IT personnel delivering the IT products and services; and the technical infrastructure, comprising the applications, data, and technology configurations. It is noteworthy that information systems are a significant constituent of the technology infrastructure, and contribute greatly to any emergent characteristic of the infrastructure.

Part of the reason for this great interest in flexibility may be found in the transformation of society and the socio-economic activities. As opposed to industrial society, the currently evolving postindustrial society is characterized by a larger, and continually increasing, amount of knowledge, complexity and turbulence. This is, in part, due to factors such as global competition, and pressure for a compressed cycle for new products (Lee and Leifer 1992). Other characteristics of this society include discontinuities, strategic management that depends on flexibility and rapid response for effectiveness and efficiency, and greater alliances, as well as sharing of information, between various organizations. There is also an increase in global strategizing that transcends national boundaries, and a corresponding increase in the diversity of stakeholders in any organizational venture (Davis 1987).

4.1.2 Characterization

4.1.2.1 Definition. A standard English definition of flexibility is " The quality or state of having a ready capability for modification or change, with resulting adaptability to new situations." (Merriam-Webster 2002). Synonyms include the following: adaptability, portability, customizability, plasticity, agility.

There are many aspects to the term flexibility. For example, in production systems, new types of equipment are said to bring flexibility to the production environment when they allow rapid changes in products without the need for a lot of new capital investment (Beyers and Lindahl 1999). By analogy, flexibility in software would allow quick and easy configuration of the end result of the manipulation of a given input without unnecessary overhead.

Flexibility can also be defined as a system's reaction to the uncertainty in its environment, and is a reflection of that system's ability to effectively adapt or respond to change. In this context, the field of supply chain management defines product flexibility as "the ability to handle difficult, non-standard orders, to meet special customer specifications, and to produce products characterized by numerous features, options, sizes and colors" (Vickery, Calantone et al. 1999). By analogy, flexibility in software can be thought of as the software's ability to handle a wide variety of input or commands, and to give output that is equally varied in nature. Similarly, just as supply chains strive to achieve flexibility to enable them to be responsive to target markets, software should be flexible in the sense of various different users being able to use it.

Heinl, Horn et al. 1999 suggest a classification scheme for flexibility in the context of workflow management. The concepts however seem widely generalizable. The following definitions apply to their work. A workflow management application performs the computer-supported execution of business processes belonging to a certain application area. Workflow types or workflow schemes model business processes. They may be instantiated in order to represent the occurrence of a business process. A workflow instance, or workflow, is the result of such an instantiation, and describes actual executions of business processes. Workflow types and the derived workflow instances together constitute a workflow management application. In this context, Flexibility by selection is defined as the multiple alternative execution paths the user has for executing a workflow. This is useful when a large number of workflow execution paths are known or anticipated.

It also has the notable requirement that it has to be anticipated, and has to be included in workflow type specification. This might however not always be possible. For example, because of the changing business process, the actual flow type may no longer adequately represent the business process. This may call for the ability to incorporate another hitherto unforeseen execution path by modifying a workflow. This is defined as flexibility by adaptation. Flexibility by adaptation requires that the workflow management system have the functionality and tools to enable the changes in the workflow type and integrate these during runtime.

In their work on the concept of flexibility in the context of information technology alignment in turbulent environments,(Knoll and Jarvenpaa 1994) define flexibility as " the ability to vary the reach of technology, to vary the range of technology features, and to use 'time' as an ally in the quest for flexibility." They further suggest three categories of flexibility as follows:

Flexibility in functionality: This refers to the capacity for input flexibility. It results in an ability to withstand variability in input.

Flexibility in use: This is the ability to recognize opportunities and exploit relationships. It is useful in a system when the outcome is unknown, and allows for proactive searches for new solutions and responses to changes in the decision maker's perception of the environment.

Flexibility in modification: This refers to the ease and variability of technology modification. Some of the factors that are relevant in modification flexibility are the type of feedback required, the type of adjustment process for bringing about this modification, and the effort required for making the adjustments.

Flexibility can also be defined as one of the desirable qualities of an organizational information system that result in the system "yielding satisficing (satisfactory and sufficing) performance under all environments that are assessed as having an appreciable probability of occurring, with an associated scanning component that alerts it to environmental changes and enables switching to another configuration quickly and inexpensively to take advantage of the environmental change" (El Sawy and Nanus 1989). From a systems perspective, flexibility can be viewed as a system's ability to 'fit' different environments and not be rendered ineffective by changing circumstances (Whitworth and Zaic 2003).

4.1.2.2 Description. The current popular focus of multimedia on video and audio applications provides a good illustration for flexibility. While these currently pose a challenge to distributed computer networks, there are far greater possibilities. For instance, rather than the computers being able to display just one medium, they should be able to "gracefully and efficiently" handle a wide variety of data flows of differing requirements, which might not even have existed when the computers or their operating systems were designed. The variety of the kind of data flows should be limited only by the creativity and imagination of the human users (Montz and Peterson 1998).

Montz and Peterson 1998 suggest that in order to achieve the desired level of flexibility that is critical for large social applications such as the world wide web, email, DNS and file systems, it is necessary to understand the flexibility demands placed on the applications, as well as the manner in which the software can be changed, both before and after it is run. Some of the demands placed on distributed systems in this regard include the following:

- They must run on a variety of hardware and operating system environments.
- The various application parts running on different machines must cooperate with each other.
- They must be able to perform vaguely defined tasks which change over time.
- Their future behavior is unpredictable as it depends on large scale human activity.
- Even though not necessarily fault tolerant or secure, their failings in these areas should be understandable and predictable.

Attributes of flexibility include efficiency in switching to other configurations as required, and the amount of variability in the environment that the system can cope with. Also, in addition to these 'reactive' qualities, this property includes preemptive capabilities that would enable it to have strategic flexibility and be able to respond in anticipation of adverse or beneficial conditions (El Sawy and Nanus 1989).

Analysis of the literature on the description of flexibility suggests three distinct aspects:

- Flexibility by detection: the system detects, recognizes and predicts a wide variety of environment input changes.
- Flexibility by response: the system has a wide variety of responses to unpredicted situation that can be quickly activated (e.g. by reconfiguring or recombining responses from its response repertoire).
- Flexibility by adaptation: whether the system can predict or reconfigure well or not, it can modify its own ability to predict and reconfigure using environment feedback.

An example of flexibility by detection or recognition is Windows "Plug 'n Play", where the operating system automatically detects physical environment changes, like adding a mouse or printer. An example of response flexibility is the preferences or options module in most software that allows users to change system operation. Adaptive flexibility corresponds to what biologists and psychologists call "learning". The realm of self-evolving dynamic systems is still a new area of IS development, but simple examples are where software "remembers" past user actions and carries them forward, such as an auto-correct feature that notices user corrections and repeats them (Whitworth and Zaic 2003).

The above aspects are distinct, and not alternatives. A system advanced in all three would be maximally flexible (in input analysis and prediction, in response variety and agility, and in adapting itself to new environments).

4.1.3 Conclusions

The flexibility construct is well documented in the IS and other literature. Various definitions are given which vary somewhat depending on the context. However, the WOSP definition of flexibility corresponds quite closely with the layman's definition given in a Standard English dictionary.

Some IS literature suggests that flexibility has several attributes, and two such suggested are efficiency in switching to other configurations as required, and the amount of environmental variability that a system can cope with. Moreover, the definitions of flexibility reviewed suggest that this construct can further be categorized into flexibility by detection, flexibility by response, and flexibility by adaptation.

4.2 Reliability

4.2.1 Significance

The increasing importance of the reliability of software today is directly linked to the wider phenomenon of increasing dependence by society on this technology. (Littlewood and Strigini 2000) enumerate several reasons for this dependence. These are summarized as follows.

Older technologies in safety or mission critical applications are being replaced with software based systems. For example, software is increasingly used for aircraft engine control functions, railroad interlocking, and protection systems for nuclear plants. Moreover, new applications that are critical in their domain have been developed. A prime example is the automation of certain surgical procedures. The common denominator among all these systems is the requirement for very high reliability. The reliable performance of the software employed in such systems is absolutely essential.

Software has increasingly been shifting from a supporting function to a decisive role in the provision of critical services. As an example, in air traffic control, software has traditionally been regarded as non-critical since human controlled manual systems could always be activated as backup. However, the increasing air traffic volumes have led to the easy overwhelming of such backups, leaving the software based methods as the only feasible way of handling the new situation.

Software is the only means of performing functions that, although they may not be perceived as critical, have great impact in their applications. Examples of this include inventory and payroll database software for such organizations as hospitals and supermarkets.

Software-provided services have become increasingly integrated with everyday life. For example, spreadsheet programs are in widespread use for such functions as decision-making.

Increasingly, software systems are integrated and interacting with each other without human facilitation. In such circumstances, software failures can propagate themselves or cause widespread disruption before there is any human intervention.

More generally, reliability has been of major concern to traditional corporate information management systems (Rinard 2003). Here, the consequences of downtime are felt directly as disrupted operations, lost profits, and dissatisfied customers who may easily shift allegiance elsewhere. As interconnectivity among software increases, more of the business functions are unable to operate absent a working system. Also, many, and increasingly more, embedded systems control critical physical operations, and their failure can have very dire consequences, making reliability an issue of paramount importance. These embedded systems are widely touted as the next software area that will experience swiftest growth (Rinard 2003).

The strength of the demand for software reliability varies with the industrial sector in which the software is employed. The two extremes in this range are generally perceived to be mass-marketed PC software on the one hand and safety-critical software such as that found in heavily regulated industries.

4.2.2 Characterization

4.2.2.1 Definition. A standard English definition of Reliability is "The quality of proven consistency in giving the same or satisfactory performance in successive use despite internal variations such as part failure" (Merriam-Webster 2002). Synonyms include the following: stability, dependability, recovery, durability, ruggedness.

Reliability can be defined as one of several attributes or properties comprising dependability of a system that give confidence in the system performing as required (Brocklehurst, Kanoun et al. 1991).

More conventionally, software reliability is defined as the probability that a software system will operate without failure for a specified time under specified operating conditions (Whittaker and Voas 2000).

Furthermore, internal robustness has also been defined as a software property of having low sensitivity to changes in the values of parameters and internal functions that may be different from the original assumptions (El Sawy and Nanus 1989).

And from a systems perspective, reliability is defined as a system's ability to recover performance despite internal variations such as part failure. (Whitworth and Zaic 2003).

A common theme underlying all the above definitions of reliability is the notion of continued performance as required despite unanticipated challenges.

4.2.2.2 Description. Reliability as conventionally defined has the notion of the duration of performance time as a factor of reliability. This may however be misleading, as can be seen by comparing software to hardware. Hardware wears over time, and may eventually fail from this wearing. In contrast, software does not change over time. Rather, it fails because of embedded defects that had so far escaped detection. Thus, time is not necessarily a factor as to when failure arising from the exposure of such a defect will occur. The definition does not take into account the complexity of the software, as well as to how thoroughly the testing is done, with ideally a wide variety of tests covering an equally varied set of operating conditions for the software (Whittaker and Voas 2000).

Similarly, the operating conditions need to be more clearly specified. The operational profile of software has been defined as comprising both the set of operations that software can execute, together with the probability that these will actually be executed. Certified reliability however applies only to the profile used in testing. Moreover, in specifying a system profile, account must be taken of more than just the primary user of the system. Most software applications have more than one user. Other users would include the operating system and other applications, which may cause the software in question to fail. Thus the operating environment of software can be extremely diverse and complex. Ideally, a specified operational profile should include all users and all operating conditions that can affect the software of interest, a feat that is as yet not possible.

El Sawy and Nanus 1989 describe a robust information system as one that is capable of the maintenance of key structural relationships when environmental disturbances occur, and which gives a system its resilience, which is the ability to absorb environmental shock and bounce back. This is in contrast to the property of flexibility and adaptation whose value is the ability to move to new equilibrium positions involving structural changes.

One of the major research challenges for the future on the reliability of software as identified by (Littlewood and Strigini 2000) is "focusing on user-centered system-level dependability qualities". Currently, it is usual to describe reliability in terms of compliance of a program with its written specifications. This has several setbacks. For example if the program is poorly written with imprecise requirements, this might mean that its reliability requirements cannot be written. A more practical approach is to define failure in terms of the effect of the system on its user. This is because users usually classify failures by their consequences rather than by their causes. An illustration, they give the example that a typical user would not care whether he loses an article he was typing because of a bug in the word processor or because the platform in use allows another application to interfere adversely with the word processor, or even because the word processor's help manual does not adequately explain a command. And indeed most users would actually be unable to tell the cause of the problem with their word processor.

The standard way of achieving reliability has been simplicity, which is keeping the program sufficiently simple that it can function with standard techniques. However, as the need for increased functionality causes a corresponding rise in complexity, this becomes an infeasible solution to software reliability (Rinard 2003). Another cause of complexity that has a negative impact on reliability is increased scalability (Clarke, Tilker et al. 2003).

Internal robustness measures are currently achieved by designing systems hierarchically with loosely coupled parts, each of which is designed with high cohesion. The methods used to institute internal robustness include structured design, dataflow diagrams, data dictionaries, database management systems using relational data models rather than hierarchical or network models, software reusability techniques, and object-oriented programming (El Sawy and Nanus 1989). By reducing the number of internal system interconnections, designers reduce the effect one part failure has on the total system. This general approach can be called reliability by modular design. Moreover, the probability of failure of a component in a system can be halved by connecting it in parallel to an identical one (Prayurachatuporn and Luigi Benedicenti 2001). This then increases reliability through redundancy. Finally, while one can design and duplicate to prevent failure, an alternative is to allow it, but recover quickly, as software error routines attempt to do (Whitworth and Zaic 2003). Hence a recommended means to reliability is the deployment of error detection techniques. No matter how well an application has been tested or designed, actual failure is possible if not probable, so error correction sub-systems are necessary. Monitoring includes not only detecting, but also logging problems for future processing.

The above analysis suggests three distinct aspects to reliability:

- Reliability by modular design: The system is designed to minimize coupling so if one part fails others still continue.
- Reliability by redundancy: The system is designed to have independent means to the same end, so if one fails another can take over.
- Reliability by recovery: The system is designed to deal with any failure by some recovery technique.

These aspects are distinct, and not alternatives. A system advanced in all three would be maximally reliable in being modular so no failure cumulated, in redundancy so no failure was critical, and in recovery so no failure was permanent. Reliable systems should continue to operate under conditions of high stress, and if they are affected, should only suffer a gradual and graceful degradation in performance, rather than crashing instantaneously. Moreover, even after failing, it should be possible to repair them within a short duration.

Much of the system design in computer systems goes into the design of software. Errors may come in at any stage of its production: requirements specification, design, implementation, testing and debugging, or in the maintenance (Borning 1987). It is however at the requirements specification level that the interaction of the system with the outside world is expressed. This implies anticipating all the circumstances for the use of the system. (Cortellessa, Singh et al. 2002) note that while early validation of functional requirements is supported by well-known approaches, that of non-functional requirements such as reliability is not.

4.2.3 Conclusions

Reliability is well documented in the IS literature as a software requirement of significant and increasing importance. Various definitions are available, some of which differ considerably in emphasis. However, the definition for reliability assumed by the WOSP model is quite consistent with the everyday meaning of the term 'reliability' as given by a Standard English dictionary.

4.3 Security

4.3.1 Significance

Increasingly, networked computer systems are becoming the site of people's work and other activities. Thus, for example, millions transact over the Internet, managing their finances and paying their bills online. Also, corporations are increasingly using the Internet to connect geographically disparate offices, or to form virtual teams and companies for a particular purpose. Furthermore, new uses for the Internet are being discussed, among them online voting systems for political elections (Dourish and Redmiles 2002). Undoubtedly, the increasing volume and diversity of information that is available over the Internet and local networks as well has brought about major improvements in the efficiency and effectiveness of diverse institutions. However, this situation has also caused increased dependency on the networked resources to such an extent that disruptions in their availability can have drastic and negative effects on the daily operations of such institutions, which include research centers, major multinational firms, universities and government departments, among others (Goan 1999). Moreover, the vast majority of these activities involve the disclosure of sensitive personal information, and many people are hesitant to do so, especially given the media highlights of such issues as credit card fraud and identity theft. More specifically, electronic commerce, which is the leading Web-based application, is projected to have a market in excess of $1 trillion over the next few years. However, information security has been a major stumbling block. Other major web-based applications with similar security concerns include telemedicine-based healthcare services and online services and businesses involving the public and private sectors (Joshi, Aref et al. 2001).

All this makes security a critical goal for information systems. Many of these security concerns have to do with the Internet's peculiar characteristic of openness and ubiquity, which make it very susceptible to various kinds of attacks. More specifically, the Internet is particularly enticing given the massive interconnection of diverse and distributed systems, the great amount of sensitive information maintained by end users such as corporations and government agencies, and the ease of distribution of automated malicious software. This last reason is all the more serious, considering the ease with which such crimes can be executed anonymously, and across geographic boundaries, and the difficulty in gathering sufficient forensic evidence of such crimes, which makes the apprehension of criminals correspondingly difficult (Joshi, Aref et al. 2001).

Some of the news with regards to breached computer security in recent times has included the penetration by hackers of the security system of America Online, the world's biggest online service provider, and their gaining access to the personal information and credit card numbers of subscribers. Others include an extortion attempt by an overseas intruder who stole a database containing personal information from a payment processing company, and posted some of the information online as threat (Rombel 2001).

It is important for all organizations today to appreciate the consequences of the breaching of the security of their information systems. These consequences may be of a financial nature, in the form of immediate costs and losses; or may be 'hidden', and of a much more serious nature. Some of the ways in which a breach might negatively impact an organization include (Farahmand, Navathe et al. 2003):

- The brand image, public standing, and marketplace goodwill.
- The financial value of business transactions.
- The confidence of the customers and general public regarding the accuracy and fraud-resistance of business transactions.
- The ability to maintain a timely revenue cash-flow.
- The ability to meet the requirements set by regulators.

Given such impacts as those above, it is important to bear in mind that a given security situation could conceivably impact two different organizations, albeit in the same industry, in completely different ways. It was estimated by the Forrester Research Group that, largely because of the Web, US business spending on security would in the year 2004 total $19 billion (Rombel 2001).

4.3.2 Characterization

4.3.2.1 Definition. A standard English definition of Security is "The quality of being shielded from hostile interference, including unauthorized entry, misuse, takeover or destruction"(Merriam-Webster 2002). Synonyms include the following: defendability, protectiveness, resistance to attack.

Computer security has been defined as "that body of technology, techniques, procedures and practices that provides the protective mechanism to assure the safety of both the systems themselves, and the information within them, and limits the access to such information solely to authorized users" (Ware 1984). More comprehensively, (Joshi, Aref et al. 2001) and (Jajodia 1996) give the following definition of security and related terms: "Information systems security refers to protection of information systems against unauthorized access to or modification of information, whether in storage, processing or transit, and against denial of service to authorized users, including measures necessary to detect, document, and counter such threats. The main goals of information security are confidentiality or secrecy, integrity, availability, accountability, and assurance. The goal of confidentiality is to ensure an unauthorized person does not access the information. The goal of information integrity is to protect information from unauthorized modification. Information availability ensures that the information is available when needed, and is not made inaccessible by malicious data-denial activities. Information accountability ensures that every action of an entity can be uniquely traced back to the entity. Security assurance is the degree of confidence in the security of the system with respect to predefined security goals."

More generally, and from a systems perspective, security can be defined as the ability of a system to protect itself from unauthorized entry, misuse, or take-over (Whitworth and Zaic 2003). If unauthorized entry, misuse, or take-over are construed as an attack against the system, this definition becomes "the ability of a system to resist attack".

4.3.2.2 Description. Security as it pertains to software and information systems can be understood by making comparisons and analogies to various aspects that have parallels with software reliability (Littlewood and Strigini 2000) . Some of these comparisons and analogies are described in summary below.

In reliability, the input space is the sum of all inputs that could possibly be encountered. A fault can be thought of as a subset of this space. Thus, when input is selected from this subset, a failure results. Similarly with security, the input space can be viewed as the set of all possible inputs to the system, and includes those involved in normal use, as well as those that are the result of intentional attacks upon the system.

The usage environment determines the mechanism by which inputs will be selected from the input space. The analogy for this in security would be all the attackers and how they behave, as well as the normal system operation. In this context, an attack would include:

- ‘passive’ attacks, as in listening to and analyzing information traffic.
- illegal use of the system, including the inserting of trap-doors by authorized system users.
- attacks on, and manipulation of, personnel, including such tactics as bribery and blackmail to force them to do something with the system they otherwise would not do.

In the context of reliability, the system boundary is conventionally drawn rather tightly, in the sense that software reliability is seen as dealing with failures of well-defined programs. This view can be misleading, when it is considered that the interaction of people with systems can be a significant source of failure in a larger picture. The inappropriateness of such a narrow view is even more apparent in the context of security. This is so as it is often because of the interaction of attackers with the computer systems and their owners that the security concerns come about.

Reliability deals with the central concept of system failure. The analogy for security is security breach. By this is meant the event where the behavior of a system deviates from the security requirements.

System failures are caused by faults in the system, where these can be specification and design faults that come about as a consequence of human inadequacy, and which in fact is the only source of failure for computer software. By analogy, vulnerabilities, which can be viewed as special types of faults, are the cause of system security breaches. Moreover, such faults can be either accidental or intentional, with the intentional faults being malicious (Trojan horses, trap-doors) or not (as for example resulting from a deliberate trade-off between security and efficiency).

Intentionality could refer to the situation whereby a system’s vulnerabilities are deliberately caused during design. It could also refer to the manner in which faults present in a system are activated during the system's operation, thus causing security breaches or failures. Whereas there is no clear distinction between reliability and security, the concept of reliability can be seen as mainly concerned with accidental failures resulting from accidental faults. Even though this might also be a reliability issue, security deals mainly with intentional attempts to cause a breach by exploiting a vulnerability that may be intentional or accidental. Furthermore, breaches, whether caused by deliberate action or not, that result from malicious intentional vulnerabilities are a security issue.

The National Institute of Standards and Technology recommends that information system security managers should adopt technologies that ensure that protection, detection and recovery mechanisms are all incorporated according to an engineering principle referred to as defense-in-depth. Given that security technologies frequently fail for various reasons, the deployment of more than one countermeasure against expected threats is recommended. The primary function of a protection mechanism is to prevent a security threat from succeeding in the first place, thus reducing threat frequencies. A detection mechanism primarily identifies that an attack is occurring, or has happened. Recovery mechanisms detect potential damage and enable system administrators to reverse the situation if possible (Butler 2002). In summary, while a protection mechanism serves to thwart the success of a threat, detection and recovery mechanisms mitigate the outcome of threats from compromised systems. All security technologies fall into at least one of the three categories of protection, detection, or recovery (Butler 2002). However, recovery, though included, is more appropriately classified under reliability. Some of the more common technologies can be classified as per Table 4.1 below, which is adapted with some changes from (Butler 2002).

Table 4.1 Security Enhancement Technologies

Protection	**Detection**	**Recovery**
Packet filter firewall	Host-based IDS	Log analysis applications
Application firewall	Network monitors	Back-up and recovery tools
Circuit firewall	Net-based IDS	Load balancing
Smart cards	Auditing	Key stroke replicator
Authentication policy servers	Key stroke replicator	Forensic software
Virtual private networks	Auditing	
Anti-virus software	Email content inspection	
Line encryption	Vulnerability assessment	
Hardened OS		

There are several perspectives from which information security and its management can be approached (Eloff and Eloff 2003). One is from a strategic perspective, which addresses corporate governance, policies, and pure management issues, among others. A second one is from a 'human' side, dealing with such issues as security culture, awareness, training and ethics. The technology perspective may focus on hardware and software issues, while a process perspective deals with the implementation of controls as per standards or codes-of-practice and the compliance of these controls. However, for an information security management system to be successful, it must incorporate all these different perspectives and be holistic in nature.

A threat is manifested by a threat agent using particular penetration techniques to cause harm in an information system. Thus, threats to information systems can be viewed from two angles: Threat agent, and Penetration technique (Farahmand, Navathe et al. 2003).

4.3.3 Conclusion

Security of software is a major concern according to the IS literature. This in part has to do with the increased dependency by society on IT in general and networked resources in particular. The openness of and ubiquity of resources such as the Internet make it particularly susceptible to various kinds of attacks. Whereas most of the definitions of security in the literature are in reasonably close agreement, there seems to be some confusion between security and privacy. Also, the definition of security adopted by WOSP closely conforms to the layman's definition of the term found in a Standard English dictionary.

4.4 Extendibility

4.4.1 Significance

Several trends in the arena of commerce have led to the increasing importance of distributed systems and accompanying scalability.

First there has been an increase in the number of mergers between companies. This has usually meant that different IT systems have to be integrated to enable the different parts of the new entity to deliver seamless service to their customers. Since the time available for such integration is usually too short to build a new system, this usually means that the existing components have to be integrated into a distributed system.

Second, due to the decreasing time available for providing new services, new systems can frequently be fashioned in a timely manner only by using components procured off the shelf and then integrated together, rather than building a system from scratch. Since such components may have different hardware and operating system requirements for example, they may be deployed on different hosts, leading to distributed systems.

Third, the Internet provides new opportunities for offering diverse products and services to a vast array of potential customers. To do this, the e-commerce sites require a level of scalability that can only be attained by distributed software architectures. Scalability identifies the dependency between the number of distributed system resources that can be used by a system, and the latency, which is the delay between request and completion of an operation, and throughput or the number of operations that can be completed in a given period of time (Denaro, Polini et al. 2004).

Over the past few years, e-commerce, which is the use of the internet to buy and sell goods and services, has experienced substantial growth. Even though this may be a new business paradigm, the underlying basic business principles for success remain the same (Arlitt, Krishnamurthy et al. 2001). Thus for example, in addition to attracting new customers, a business must strive to retain its existing customers, raising profits through extended customer relationships. It has been shown that retaining 5% more customers results in increase in profits of up to 100% (Arlitt, Krishnamurthy et al. 2001). One of the key conditions for retention of customers is providing a pleasurable shopping experience. This, among others, includes ease and promptness in finding goods or services, and minimum delays for making a purchase. If a web based shopping system capacity is inadequate, system performance can degrade dramatically with increase in customers as a result of a business promotion effort, for example. Similarly, whereas features such as personalization of service are very important for building customer relationship, they typically take a toll on the capacity of the system (Arlitt, Krishnamurthy et al. 2001). This would in all probability lead to a loss of customers because of poor service. To avoid this, the capacity management of such systems should be governed by scalability. Specifically, the systems should have adequate horizontal scalability, which means that as the number of concurrent users of the system increases, more servers are added.

On a slightly different perspective, nearly every type of business has experienced an explosion in the amount of data that it must store, process, and understand. Simultaneously, the acceptable time frame for processing all of this data into information continues to grow shorter. The organizations that are in the best position to handle this situation are those that are able to adjust or scale their hardware and applications to accommodate this increased demand. Thus scalability of a system subject to growing demand is critical to its long-term success.

As a specific example, scalability has been identified as an important requirement for workflow management systems. These are networked computer systems that are designed to help organizations coordinate, monitor and execute their various activities within a distributed work environment. They must be able to deal with issues such as internet-based large scale applications, dynamic change and ad-hoc work patterns, as well as other social aspects of the work. Thus, performance requirements include flexibility, robustness, modifiability, availability, and usability. And as these systems become more complex, scalability is becoming an increasingly critical performance criterion (Kim and Ellis 2001). Two aspects of scalability have been identified as critical for these systems (Kim and Ellis 2001) as discussed below.

The system must be scalable to the varying workloads. Thus it should provide approximately the same response time regardless of the number of active users and concurrently running workflow instances.

The system must be scalable to the size and structure of the enterprise within which it is deployed. Thus for a global enterprise, the workflow management system must integrate many semi-autonomous and even fully autonomous organizational entities such as branches and departments in order to implement enterprise-wide business processes.

One advantage of software extendibility would be decreasing the need to write software, as that already existing can have increased functionality by having it work together with other applications (Meyer 1987).

4.4.2 Characterization

4.4.2.1 Definition. A standard English definition of Extendibility is: "The capability of having expanded functionality with the addition of a relatively inexpensive or otherwise readily available application or tool so as to reduce cost, improve efficiency, or attain some other desired effects" (Merriam-Webster 2002). Synonyms include the following terms: openness, compatibility, scalability, open standards.

In its most general form, scalability has been defined as "The ability of a solution to a problem to work when the size of the problem increases" in the dictionary of computing at http://wombat.doc.ic.ac.uk (Rana and Stout 2000).

Alternative definitions of scalability are given as follows in the link: http://search390.techtarget.com/sDefinition/0,sid10_gci212940,00.html:

- The ability of a computer application or product (hardware or software) to continue to function well when it (or its context) is changed in size or volume in order to meet a user need. Typically, the rescaling is to a larger size or volume. The rescaling can be of the product itself (for example, a line of computer systems of different sizes in terms of storage, RAM, and so forth) or in the scalable object's movement to a new context (for example, a new operating system).

- The ability not only to function well in the rescaled situation, but to actually take full advantage of it. For example, an application program would be scalable if it could be moved from a smaller to a larger operating system and take full advantage of the larger operating system in terms of performance (user response time and so forth) and the larger number of users that could be handled.

Mainly from a design perspective, four types of scalability can be defined: load scalability, space scalability, space-time scalability, and structural scalability (Bondi 2000):

- Load scalability refers to a system's ability to function gracefully, or without undue delay and without unproductive consumption of resources at light, moderate and heavy loads while making good use of available resources.

- Space scalability: A system or application is said to have space scalability if its memory requirements do not become insupportable as the number of items or objects it supports increases. This is subjective, but for example, an application would be said to be space scalable if the memory requirements increase at most sublinearly with the increase in the number of items.

- Space-time scalability: a system or application has space-time scalability if it continues functioning gracefully as the number of objects it encompasses increases by orders of magnitude. This can be achieved if the data structures and algorithms used to implement it are conducive to smooth and speedy operation regardless of system size. For example, a search engine based on linear search would not be space-time scalable, while one based on an indexed and sorted data structure such as a hash table or balanced tree could be.

- Structural scalability: A system is said to be structurally scalable if its implementation or standards do not impede the growth of the number of objects it encompasses, at least not within a given time period. This is a relative term since scalability depends on the number of objects of interest now relative to the number of objects later.

In addition to the above aspects of scalability, two others are defined:

- Distance scalability: this refers to and an algorithm or protocol working well over long distances as well as short distances
- Speed/Distance scalability: here, the algorithm or protocol works well over long or short distances, at high and low speeds

The online encyclopedia, Wikipedia, gives the following definitions of various aspects of scalability:

In telecommunications and software engineering, scalability indicates a system's ability to maintain quality performance or service under an increased system load by adding resources (usually hardware). A system whose performance improves after adding hardware (proportional to the capability added) is said to be a scalable system.

In general systems, scalability is the measure of a system's ability to increase or decrease in performance and cost in response to changes in application and system processing demands. Examples would include how well a hardware system performs when the number of users is increased, how well a database withstands growing numbers of queries, or how well an operating system performs on different classes of hardware. Enterprises that are growing rapidly should pay special attention to scalability when evaluating hardware and software.

In hardware, scalability is the ability to increase the size and processing power of an online transaction processing system by adding processors and devices to a system, systems to a network, and so on, and to do so easily and transparently without bringing systems down.

In a database, scalability is when performance remains unchanged despite the number of queries or transactions, or a server performance remains unchanged despite the number of ports or users connected.

In routers, scalability is the capability of a product (hardware or software) to function well as it scales up or down to meet a user's needs. For example, a router (hardware) with one wide area network (WAN) port is said to be scalable if it can be scaled up to support more than one WAN port or down (back to one WAN port) through the addition or removal of WAN modules.

In routing protocols, the scalability of a method measures its capability to perform efficiently as parameters of the network increase to large values. As a typical example, a routing protocol is considered scalable with respect to network size, if the size of the necessary routing table on each node grows as O(log N), where N is the number of nodes in the network.

To scale vertically or scale up means to add resources to a single node in a system, such as adding memory or a faster hard-drive to a computer.

To scale horizontally or scale out means to add more nodes to a system, such as adding a new computer to a clustered software application.

From a system's perspective, scalability can also be defined as a system's ability to incorporate or use outside elements in system operation (Whitworth and Zaic 2003).

4.4.2.2 Description of Extendibility. Scalability is a desirable attribute of any network, system, or process, and connotes the ability of a system to accommodate increasing numbers of elements or objects, to process growing volumes of work gracefully, and/or to be amenable to enlargement. Hence the importance of this quality in any system to the extent that it is often specified in procurement contracts (Bondi 2000). On the other hand, "When we say that a system is unscalable, we usually mean that the additional cost of coping with a given increase in traffic or size is excessive, or that it cannot cope with this increased level at all. Cost may be quantified in many ways, including but not limited to response time, processing overhead, space, memory, or even money. A system that does not scale well adds to labor costs or harms the quality of service, and can deprive the user of revenue opportunities. Eventually, it must be replaced." (Bondi 2000).

It is widely acknowledged that scalability is not a characteristic of hardware or software exclusively, but of both (Law 1998). To deal with this, most researchers keep the hardware properties constant in order to evaluate the software, or conversely, they keep the software properties constant while evaluating the hardware on scalability.

To illustrate software scalability, (Whitworth and Zaic 2003) give the example of viruses which are software systems with little functionality on their own, but which are able to take over other systems and use them to great effect.

In order to be extendable, it is necessary for software systems and components to match. One way of achieving this is to have standards that the systems and components must adhere to. Another way is for a given system to avail its extension points (Whitworth and Zaic 2003).

4.4.3 Conclusions

A major advantage of extendibility in software is that it decreases the need to write more software as that already existing can have its functionality enhanced by having it work with other applications. Other frequently used terms in the literature are scalability and openness. Most of the definitions of this factor in the literature seem to be influenced by the context or the application type in question. However, the definition of extendibility given by the WOSP model matches quite closely the layman's definition given in a Standard English dictionary.

4.5 Privacy

4.5.1 Significance

Privacy is a prerequisite in the information society for at least two major reasons. First, it provides the means to resist the commoditization of human beings, and, secondly, it is a means of enabling e-commerce and electronic service delivery (Clarke 1999).

The threat of invasion of privacy by computer and information technologies comes in myriad forms. These range from online activities such as spam, cookies and the clickstream, to real-time person location technologies such as intelligent transportation systems, geo-location, and biometric identification, among others. Also included in this category are miniature processors that are embedded in such diverse places as plastic cards, anklets, watches, rings, products and product packaging, and people (Clarke 1999). These concerns about privacy have led to a significant decrease in the trust people have in the information society. For instance, according to an early estimate by the Boston Consulting Group, as much as $6 billion in additional electronic commerce could have been gained by the year 2000 with the implementation of strong user privacy controls (Wright and Kakalik 1997). In addition to fears that hackers might steal private information during transmission to or from Web servers, most Internet consumers feel that they lack control over what Web based merchants do with their data after a transaction. Given that the success and profitability of web-based enterprises is largely dependent on both acquiring new customers and maintaining the current ones, consumer confidence is of paramount importance for their survival (Wright and Kakalik 1997).

Information linked to individual persons was formerly difficult to come by, and even more onerous to cross-reference. Today, it is much easier to collect and manipulate such data using automated search facilities. Moreover, in the past, private information, even that available through public documents such as real estate deeds and SEC disclosure forms was not readily available for mass distribution, but rather was usually locked away in physically secure places and in dispersed locations. Now, most of this information is placed in databases that are often accessible on-line. All this poses a major threat to privacy (Wright and Kakalik 1997).

Commercial and governmental use of database technologies poses a significant threat to individual privacy as the following examples illustrate (Wright and Kakalik 1997):

- Equifax, one of the largest credit bureaus in the United States, avails the credit records of more than 160 million consumers to over 50,000 businesses.

- The Medical Information Bureau, which is a centralized collection of agency for medical data, avails the medical records of over 15 million individuals to over 750 insurance companies.

- The Federal Parent Locator Service database, containing a record of every driver's license and all professional licenses issued in the United States; The Federal Case Registry of Child Support Orders database containing records of every marriage, divorce and paternity determination conducted in the United States; and the National New Hire Directory database containing the names, addresses, job descriptions, and employers' identities for all new hires in the United States, are all federal databases that are cross-indexed using the social security number.

- Internet Service Providers (ISP) and websites now collect and sell data about unsuspecting on-line users at unprecedented rates using data-gathering tools.

- Cookies are employed prevalently by on-line advertising organizations to track user activities on the Internet without the users' knowledge, enabling these agencies to deliver targeted on-line advertisement.

- Whereas technology has facilitated the collection and distribution of vast amounts of information, it does not necessarily ensure the accuracy of that information. There have been for instance cases of mistaken identity involving individuals with similar names and abbreviations.

Other privacy concerns include:

- Activities of junk email marketers.

- Malicious programs that can take advantage of security loopholes in many internet tools such as java, ActiveX, JavaScript and others to obtain individuals' credit information and access personal files.

- The distribution of private information for purposes other than what it was collected for, for financial gain.

In particular, information privacy has been under increasing threat resulting from the rapid replacement of costly physical surveillance with the systematic use of personal data systems for the investigation or monitoring of the actions and communications of people. For example, data trials of individuals provide the means for public and private organizations to exercise power over them. As an illustration, profile data can be combined with sender-driven technologies to push customized information to targeted individuals, thus exercising significant influence over their behavior and reducing their freedom to think and act independently. Moreover, data is increasingly being collected and personalized as the cost of doing so continues to decrease rapidly. Storage technology ensures the continued availability of the data, while database technologies make it discoverable, and telecommunications facilitate its rapid propagation.

Attempts by authorities to justify the low levels of privacy on the Internet include the prevention of fraud on public revenue, credit guarantors, and insurance companies;

the efficient gathering of taxes; and the efficient marketing of goods and services. In addition, law enforcement and national security have been cited as additional reasons. Given the current insufficient oversight of privacy protection, the resulting severe lack of trust by consumers has led to a much slower than expected adoption of e-commerce (Clarke 1999). This notwithstanding, Internet shopping has brought with it unparalleled convenience. Using computers, consumers can now shop for and view a wide range of products from manufacturers and retailers from almost any part of the world, all from the convenience of their homes or offices. It is thus no wonder that, as early as 1997, the Internet was estimated to reach 50 million people worldwide, with monthly growth rates averaging 10%. Nevertheless, a survey found that two thirds of the respondents considered privacy concerns to be very important (Wang, Lee et al. 1998).

In the realm of databases, data mining and knowledge discovery are two new research areas that have benefited greatly from advances in data collection and data dissemination technologies. These areas deal with the automatic extraction of patterns from large quantities of data. However, it is now well documented that these activities can lead to serious threats to privacy. For example, it is possible for confidential information to be derived from data, leading to what is referred to as the "database inference" problem (Verykios, Bertino et al. 2004). More generally, while knowledge discovery and data mining are powerful technologies capable of great benefit, they can equally be abused deliberately or inadvertently. For instance, given that data mining is an inherently inductive process, many of the rules may be overly generalized, leading to sociological stereotyping, or may, due to inadequate statistical analysis, appear useful, but may actually be misleading. This could lead to negative stereo-typing and invasion of privacy for the subjects of the rule, and loss of reputation and litigation for the publisher of the results. There is therefore a need to evaluate the rules in terms of its consequences on privacy (Fule and Roddick 2004).

4.5.2 Characterization

4.5.2.1 Definition. A standard English definition of Privacy is "The quality of being able to privately or secretly communicate, convey, and act on data or information so that it is known only to a select few, rather than publicly"(Merriam-Webster 2002). The following are some of the synonyms for the term privacy: confidentiality, secrecy, camouflage, stealth, opaqueness.

Privacy can be applied in a moral or legal context, but is more understandably perceived as the interest individuals have in maintaining a personal space that is free from interference from others. Aspects of personal space or privacy include privacy of the person in the sense of the integrity of the person's body, privacy of personal behavior, privacy of personal communications, privacy of personal data, and information privacy. This latter embodies the notion that data about individuals should not be available to other persons and organizations, and that where such data is in the possession of other parties, the individual is able to exercise significant control over the data and its use (Clarke 1999).

On a different perspective, privacy can be defined as the right to be left alone, and is related to solitude, secrecy and autonomy (Wang, Lee et al. 1998). In the context of the electronic marketplace and consumer activities therein, privacy refers to personal information, with the unauthorized collection, disclosure or other use of such information in e-commerce transactions constituting an invasion of privacy.

Privacy can also be taken to mean that the subject/owner of the information can control it. In contrast, confidentiality refers to secrecy in the sense that only the intended recipients of a message read it (Araujo and Araujo 2003).

Privacy can alternatively be defined as an individual's desire and ability to keep certain information about themselves hidden from others. (Fule and Roddick 2004). However, given that participation in any society necessarily means communication and negotiation, absolute privacy in a society is not attainable (Gavison 1984).

The WOSP model defines privacy from a systems perspective as the ability of a system to control the release information about itself.

An analysis of the above yields four aspects of privacy violation, namely the collection, use and distribution of private information, and unwanted solicitation. The collection aspect includes such activities as improper electronic access to an individual's information, as for example infiltrating the personal computer of an individual while he is connected to the Internet without his notice or acknowledgement. This could result in exposing private information to unauthorized viewers. Collection also includes improper monitoring or surveillance on a software user. An example is the monitoring of an individual's activities on the Internet using such technology as cookies.

Use of private information would include such activities as merging together private information of individuals from diverse sources, and carrying out analysis on an individual's private information without the individual's notice or consent, and deriving conclusions from such analysis. Such analysis might include the individual's shopping and spending patterns, shopping behaviors and preferences. It also exposes the owner of the information to inaccurate stereotyping.

Distribution involves such activities as transfer of a customer's information to a third party without the individual's notice or consent. As an example, many internet companies sell, publish or share their customer databases containing private information.

Unwanted solicitation involves the transmission of information to individuals without their prior knowledge or permission, and using includes such methods as junk mail, and mass direct emails.

4.5.2.2 Description. Personal information can be placed in two general categories: Static private information is personal information that is not expected to change significantly over time. Examples include historical financial information, health information, personal beliefs and affiliations, and personal documents. Dynamic personal information is private information that changes significantly with time, but which can nevertheless be collected and analyzed in a manner that would result in a fairly detailed individual profile. Examples include an individual's activity history and activity content on the web (Wang, Lee et al. 1998).

Internet user privacy concerns can be categorized under improper acquisition, improper use, privacy invasion, and improper storage (Wang, Lee et al. 1998). Improper acquisition can further be categorized as follows:

Improper access: This refers to the infiltration of an internet user's private computer without his notice or acknowledgement. It results in the exposure and collection of private information to unauthorized persons.

Improper collection: This means the collection of a consumer's private information from the Internet without his notice and acknowledgement. Such information includes the email address, types of software used, the user's web access history, and private databases and files. Improper collection usually leads to improper analysis and transfer of information.

Improper monitoring: This means conducting surveillance on a consumer's internet activities without his notice or acknowledgement. It involves the use of such technology as cookies to record where, when, and for how long the individual visits web sites, and the transactions he carries out. It usually leads to improper analysis and transfer of private information.

Improper use can further be broken down into the following categories:

Improper analysis: This means the analysis of an individual's private information without his notice or acknowledgement and deriving conclusions from such analysis. These conclusions could be a consumer's shopping behaviors and preferences, and his shopping and spending patterns.

Improper transfer: This refers to the transfer of a user's private information to other organizations without his notice or knowledge. Examples of this include the selling, publishing, distributing and sharing of customer databases by some internet companies

Privacy invasion or unwanted solicitation: This is the transmission of information to potential internet users without their acknowledgement or permission. Examples include junk mail, and mass direct email.

Improper storage: This is the keeping of private information in a non-secure manner resulting in a lack of trustworthiness of the stored information, or lack of authentication control for information access. Examples include the viewing of the private information of other users by an individual user, and the changing of private information without the proper authorization. Improper storage is usually related to concerns about information confidentiality and data integrity.

4.5.3 Conclusions

Privacy is increasingly being seen as a prerequisite in an information society. Its significance as an application requirement is therefore well documented in the literature. While the existing definitions of the term have varying shades of meaning, they all have in common the notion of information about individuals only being available to other parties of his or her choice. The definition of privacy used in the WOSP model, while quite specific, fits well in the broader layman's definition of the term. The most effective means of ensuring privacy appears to be for individuals to be able to exert greater control on the collection, use, and disclosure of their personal information.

4.6 Connectivity

4.6.1 Significance

A number of studies have strongly challenged the fact that the user concept in IS research excludes context from a theoretical perspective. These studies build on the early work of human relations and socio-technical design researchers, and examine ways in which ICT designs that are based on the user concept may be inadequate, and may dehumanize, or disrupt cohesive and productive work contexts (Lamb and Kling 2003).

The world wide web brought with it a new perspective to software performance, namely social interaction. Social interaction involves communities and networks, and is an integral part of human biological behavior, given that humans are social beings. It is a performance multiplier, as societies create benefits from the synergy that comes with cooperation through communication. Communication requires two or more systems, and so systems should be designed to enable the exchange of information and meaning (Whitworth and Zaic 2003).

The significance of connectivity can also be illustrated by concrete examples. One such example of the application of enhanced connectivity is in the use of ad hoc wireless networks in a modern battlefield environment (Gray 2000). Although some locations may usually have cell phone, under battle conditions these may be destroyed or otherwise rendered non-operational or inaccessible. For such reasons, the military is investigating the use of communication mechanisms that fighters can carry with them such as ad hoc networks for data communications. In this setup, soldiers would have transmission equipment that would make each one of them a node in the ad hoc network. This is also possible with vehicles, and in fact the US military is currently using this method in field operations (Gray, Kotz et al. 2004). Likewise, ad hoc networks can be used to expedite response in civil emergencies such as remote triage for casualties (Wendelken, McGrath et al. 2003).

Similarly, the significance of connectivity is illustrated by the observed concern that the IEEE 802.11 Medium Access Control protocol for wireless networks currently does not consider the actual requirements of flows competing for limited wireless bandwidth when scheduling them (Shah, Chen et al. 2005). This results in some flows getting much more throughput than that minimally required, while others get less that what is needed to satisfy their minimum requirements. Such unbalanced allocation of bandwidth is an aspect of inefficient connectivity, and results in reduced quality of service.

More generally, some of the important aspects of connectivity in asynchronous transfer mode-based wireless networks, particularly when performing rerouting, include:

- Limiting handoff latency.
- Maintaining an efficient route.
- Limiting disruption of continuous media traffic.

Limiting handoffs is particularly critical in microcellular networks because otherwise users may frequently lose contact with a previous wireless access point. Maintaining an efficient route could also lead to disruptions in user traffic that are unacceptable for continuous media applications like audio and video that use data packets for transmission. It therefore becomes important to strike a balance between maintaining an efficient route and limiting disruptions in continuous media traffic, while keeping the handoff latency low (Ramjee and La Porta 1998).

4.6.2 Characterization

4.6.2.1 Definition. Connectivity has been defined as "The ability of a system to support information and meaning exchange about the external world, personal states and group action" (Whitworth and Zaic 2003). They give the following terms as synonyms for connectivity: interactivity, sociability, connectedness, networking, communication.

4.6.2.2 Description. Connectivity involves the communication between two or more parties in the form of passing messages. The recipient then attaches meaning to these messages. Connected software allows users to engage socially with others, or it can itself remain current by downloading updates (Whitworth and Zaic 2003). Connectivity includes how signals are routed as for example circuit vs. packet switching, and also involves such aspects as number of channels, channel bandwidth, immediacy symbol variety, and rehearsability. In order for communication to occur, there needs to be an environment or medium through which the interaction occurs, or which enables the transmission of messages. The technology in use provides this environment. So for software, this could be e-mail or computer mediated communication, analogous to telephone or face-to-face communication. However, "meaning" from an interaction is not a property of the medium, but rather the result of the cognitive process of encoding and decoding the signals that travel the medium.

Streams of symbols that are processed in the same general manner constitute a channel. And for this to occur there must exist a common processing context between senders and recipients of messages. Thus, meaning is transmitted by the entire interaction that is comprised of the sender, the recipient, the signal, and the medium. Interaction properties can be defined according to how symbols are created and received, the symbol context, the variety of the channel, the numbers of senders and receivers, how these roles change, the effort required to send and receive messages, and the ability to connect and communicate. The interaction properties can be defined as below.

Immediacy: This is the extent to which the creation and transmission of symbols is seen as immediate, continuous, irreversible, and beyond the control of the sender. Immediate communication comes directly from the sender, and is usually continuous. Immediacy is spontaneous and better reveals the sender. An example is face-to-face communication. A computer based example is email, which, even though it allows editing, comes directly from the sender, giving the sense of spontaneity and informality (Lea 1991). In contrast, rehearsibility is the extent to which a message is edited, censored, or otherwise modified prior to transmission. This gives a sense of lack of spontaneity or lack of genuineness (Dennis and Valacich 1999).

Asynchrony: This is the degree to which the decoding of symbols is seen as not immediate, continuous and irreversible, but under receiver control. It is conventionally defined as the delay between the sending of signal and its receipt, while synchrony is when message transmission time is low, so that messages are sent and received almost simultaneously (Turban 1995). However, in asynchronous interactions, the asynchrony is between the receiver and the medium, rather than between the sender and the receiver. A major benefit of asynchronous communication is that receivers can process it at any time, and can sometimes analyze and replay a message. This is because the arrival of the message does not drive the receiver's processing. The contrast is ephemerality, where signals must be processed on arrival or not at all. Synchrony, on the other hand, requires that both the sender and the receiver be synchronous with the medium of communication.

Declarative: The extent to which symbols are perceived to reference other symbols, as opposed to indicating meaning directly. Declarative communications use arbitrary symbols that must be learned in order to reference concepts, which in turn refer to actual meanings. Examples are language, mathematics, and music notation. This is in contrast to, for example, a smile, which communicates its meaning directly. A declarative version of this communicating the same meaning would be the words "I am happy". Similarly, in AI, declared knowledge exists in explicit or declarative form, rather than in the direct processing of a program (Hofstadter 1999). Since declarative symbols represent other symbols and concepts, they can be connected to each other. For example, the alphabet is very versatile, and allows an almost infinite variety of word symbols to be formed out of it. It thus overcomes the restrictions to details that pictographic languages such as hieroglyphics have. Declarative symbols such as words can represent any concept or precept, thereby allowing complex detail to be easily stored in computer databases. A major disadvantage of declarative communication is that it is at least once removed from the direct meaning. Moreover, not all communication is declarative or linguistic, and direct symbols such as facial expressions reference non-conceptual meanings directly (Maturana and Varela 1998).

Expressiveness: refers to the degree to which channels are perceived to activate available participant processing. This definition is a more general form of another one of expressiveness as allowing intonation, facial expression and gestures as well as content (Kraut, Galegher et al. 1992). Expressive communications make use of various different processing channels for the transfer of meaning, so as to evoke a more complete response. Because humans have multiple senses, and many ways of processing within each sense, to process just through one sense gives an incomplete experience. Expressive communications are more real, as they are closer to the actual perceptions people are used to.

Interactivity: This is the average perceived rate of change of sender/receiver roles for related communications or messages (Rice 1987). This is a generalization for quickness of feedback (Kraut, Galegher et al. 1992; Dennis and Valacich 1999). If averaged out over the participants of an interactivity session, it is equivalent to reciprocity or equality of participation (Rice 1994). Communication that is one-way has no interactivity or feedback, and minimal equality. Interactivity is reduced if message transmission is slow, or responder is slow or responds at length. Interactivity enables a receiver to infer that the person they are communicating with is co-present by sending a message and observing the related response. The contrast of interactivity is restricted communication.

Multi-linkage: This is the degree to which communications are perceived to represent meaning from many senders, to broadcast meaning to many receivers, or both simultaneously. It thus involves many participants at the same time. The two aspects of linkage are the number of receivers, or the audience that is broadcast to and the number of senders represented by the signal. Meaning exchange can occur on one-to-one, one-to-many (broadcast), many-to-one (merged), or many-to-many basis. On a media level, this property is referred to as network architecture (Rice 1987).

Identifiability: This is the extent to which communications are perceived to allow access to the sender's communication history. Identifiable communications allow sender history information to be carried forward or otherwise accessed. It could be in the form of a unique name that gives access to a full history, or unique voice patterns that link the same sender to his other communications. The contrast of identifiability is anonymity, which prevents any past information from carrying forward.

Cost: This is the perceived psychological effort or cost to send and/or receive information, or to interact. It includes the cost of preparing, addressing, filing, and receiving a message. "Messaging threshold" is the cost that a user finds acceptable for sending a particular message. If the cost of sending a message is greater than the messaging threshold, it is not sent. Only those messages whose urgency exceed the threshold are sent. Face-to-face involves the least cognitive effort, and is the most natural because people are adapted to it (Kock 2001). Furthermore, face-to-face messages are more quickly sent and received than any other messages.

4.6.3 Conclusions

Connectivity of information systems is increasingly been acknowledged as being important in a wide range of circumstances, especially given the increased networking of resources. Of particular prominence in this regard is the Internet. Among others, the importance of connectivity has to do with the efficiency of service of an application, particularly in ensuring adequate throughput of data. Connectivity is as yet relatively new in IS literature, and there are not many concrete definitions of the term.

4.7 Functionality

4.7.1 Significance

The need for functionality is self-evident, and involves the provision of functions required for the performance of given tasks. Thus the specification of the requirements for a system's functionality involves the obtaining from potential users of the list of desired functions, which the designers proceed to furnish. Comparing the initial users' list to the functions actually provided in the system indicates how functional the system is. Functionality is task oriented as it is relative to the task that the user wants to accomplish with the system (Goodwin 1987).

The advantage of functionality is getting the job done, or effectiveness (Whitworth and Zaic 2003).

4.7.2 Characterization

4.7.2.1 Definition. A standard English definition of 'functionality' is "The quality of being useful in the context of a given purpose" (Merriam-Webster 2002). Synonyms include: effectiveness, capability, usefulness, power, strength.

Perceived usefulness has been defined as the degree to which a person believes that using a particular system will enhance his or her job performance (Davis 1989). From a systems perspective, functionality is a system's ability to act upon its environment. Distinct from just what the system does, this refers to what it does to its' environment (Whitworth and Zaic 2003).

4.7.2.2 Description. From a user's perspective, a system has a high perceived usefulness if the user believes that there is a positive use-performance relationship (Davis 1989). A system's functionality can be described as a set of environmental changes that it can effect (Whitworth and Zaic 2003). The greater the number of changes it can effect, then the greater the functionality. In the case of an information system, functionality is the ability to change information, as, for example, a word processor's ability to change a word document. However, an emphasis on functionality could lead to software that has so many features that the user has difficulty using it (Whitworth and Zaic 2003).

The functionality of a system can be grouped under the following general headings (Sommerville 2000):

- Sensor components: These components gather information from the system's environment. Examples include the radars in an air traffic control system, paper position sensors in a laser printer, and the thermocouple used in a furnace. An example in software systems is input statements.

- Actuator components: These cause change in the environment of the system. Examples of actuators include valves that control the rate of flow of a liquid in a pipe by opening or closing, and the paper feed mechanism on a laser printer that moves the paper across the scanning beam. An example in a software system is output statements, and assignment operations.

- Computation components: These are the components that carry out computations on a given input to produce some output. An example is a floating-point processor that does computations on real numbers.

- Communication components: These are components whose function is to facilitate the communication of the other components with each other. An example of a communication component is an Ethernet, which links different computers in a building. In software systems, examples include global variables, data structures, arguments, sockets, TCP/IP protocols.

- Coordination components: These are the components whose function is to coordinate the operation of the other components. An example of this is a scheduler in a real-time system. This decides when the different processes should be scheduled to run on a processor. Other examples in software systems include control structures, procedure call, semaphores, and a state machine.

- Interface components: These components form the interface between other system components. Thus, they transform the representation used by one system component into the representation used by another. One example is a human interface component that takes a system model and displays it for the human operator. Another example is an analog-digital converter the converts an analog input into a digital output.

4.7.3 Conclusions

The importance of functionality is so self-evident or obvious as to be hard to describe. Functionality is application specific. The layman's definition of functionality and that adopted by the WOSP model, while both quite broad and generally applicable seem to emphasize different perspectives. The dictionary definition refers to functionality as usefulness in a given context, while the WOSP model looks at functionality in the context of an application changing its environment.

4.8 Usability

4.8.1 Significance

Developments in technology and in the business environment have dramatically increased the need for improved usability in software. While systems bring technology to the users, it is important to ensure that the systems succeed in bringing information as well (Blignaut 2004). One such source of pressure is the opportunities created by advances in technology for rapid and effective presentation of forever increasing amounts of information. For example:

- Increased use of networked and distributed systems across enterprises has led to access to greater amounts of information.
- Increasing numbers of people have access to computers. Surveys indicate that about 70% of Windows users use five or more programs simultaneously, and that many use more ten or more different programs in a day.
- Greater memory and faster processing is increasingly available at ever decreasing costs, spurring the use of software, and thus bringing issues of usability into focus.

Businesses are under increasing pressure to deliver better customer service. Among others, this means that an enterprise's employees must be able to easily respond to customers' needs and that customers or users must be able to easily interact with any system the enterprise puts for his use. Users are increasingly demanding from computers power without complication (Singh and Kotzé 2002). Such demands call for well-designed user interfaces. It is easy to see therefore that a substantial bottom line impact could result from paying attention to usability by a user-centered design, through such aspects as increased efficiency, reduced training time, reduced system maintenance costs after implementation, and fuller utilization of system functionality. The converse is also true, that failure to take account of usability issues in software could have a negative impact on the bottom line.

A further cause for the increased interest in usability is the changes in the relative cost of technology versus human capital. While the cost of technology is decreasing, that of human capital is increasing. Therefore, investments that would lead to reduction of the latter in the form of training, time to complete tasks, and user errors, for example, would likely lead to improved profits for businesses. To underscore this, one survey of 6000 computer users found an average of 5.1 hours a week wasted trying to use computers. This means that more time is wasted in front of computers than on highways (Shneiderman 2000). Moreover, other recent studies indicate that even experienced users of common personal computers have approximately 45% of their time wasted with frustrating experiences tied to confusing menus, indecipherable dialog boxes and functions that are hard to find (Shneiderman 2000).

Whereas developers typically view a system's functionality as separate from the user interface, to users, the interface is the system. They usually do not make distinctions between the underlying functionality and the way it is presented in the user interface. Thus, if the interface is usable, they perceive the entire system as being usable, and vice versa. Given that the user interface is the user's access to the system's functionality, poorly designed interfaces could be a major limiting factor to companies maximizing on the benefits of the technology of their systems. To see that this is so, one needs to appreciate that the user interface is comprised of everything that the user experiences, sees and does with the computer systems, which includes:

- The match with the user's tasks.
- The metaphors used.
- The controls used.
- Navigation within and between screens.
- Integration between various applications.

Given the heightened interest and practical implications of usability, user interface represents a significant and increasing proportion of the investment in IT. The benefits of good usability in software include the following:

- Reduced errors.
- Lower support costs.
- Lower training cost.
- Less loss of productivity on introduction of a new system.
- More focus on the task being carried out, rather than on the technology.
- Better employee morale, and lower turnover.
- Reduced rework of software to meet user requirements.
- Higher cross transfer of skill among applications, which reduces the need for training.
- A fuller utilization of functionality.
- A higher quality of service.
- Higher customer satisfaction.

According to the U.S. Department of Health and Human Services, citing studies by Forrester Research, two of the most striking costs of a bad site design with poor usability are:

- Loss of approximately 50% of the potential sales from a site because of people not being able to find what they are looking for.
- Loss of repeat visits from 40% of the users who do not return to a site when they had a bad experience during their first visit.

On a different perspective, there has been renewed interest in electronic based voting systems in the United States, especially given the controversies surrounding the 2000 presidential elections. Municipalities have been investing heavily in electronic voting systems. While such systems offer the promise of faster and more accurate voting, they currently have a lot of usability problems. These systems pose a unique challenge to interface designers given the fact that unlike almost any other system, voting systems must be usable by every citizen 18 years and older. This includes the elderly, disabled, and the insufficiently literate. To compound the situation, there is little or no opportunity to train the voters.

A related but wider issue to conventional usability is universal usability. Universal usability can be defined as having more than 90% of all households as successful users of information and communications services at least once a week (Shneiderman 2000). Attaining this goal is key to meeting The Association of Computing Machinery (ACM) code of ethics which states that "in a fair society, all individuals would have equal opportunity to participate in, or benefit from, the use of computer resources regardless of sex, religion, age, disability, national origin or other such similar factors". There has been a growing interest in universal usability for information technologies spanning continents and disciplines. The motivation includes the legal requirements for providing access to disabled users, the market opportunities of reaching diverse consumers, opportunities by way of medical informatics for health support groups and healthcare information, and the practical necessities of serving people who use different technologies with diverse configurations and with diverse configurations (Shneiderman 2003).

Usability can have significant measurable effects especially for high-volume, structured tasks. For example, for computer systems used for large-scale transactions, relatively small improvements in usability can lead to savings of time with consequent huge cost savings. As an illustration, a saving of one second per transaction could translate to savings of thousands of dollars, as well as greatly enhanced productivity. Reduced task completion times and error rates are typical gains that could be brought about by improved usability (Goodwin 1987).

It has been suggested that providing for usability in a system is as important as providing for functionality. A lack of usability could lead to system failure. A system with poor usability will, at the very least, cost its users both time and money. At worst, it may not be used at all since the usefulness of its functions has not been demonstrated (Goodwin 1987).

4.8.2 Characterization

4.8.2.1 Definition. A standard English definition of Usability is "The quality of being convenient and practical for use" (Merriam-Webster 2002). Synonyms of usability include: ease of use, parsimony, efficiency, user friendliness.

Quoting from (Agarwal and Venkatesh 2002), "Usability has been conceptually defined and operationally measured in multiple ways. (Gray and Salzman 1998) succinctly summarize the state of affairs related to the definition of usability noting that "The most important issue facing usability researchers and practitioners alike is the construct of usability itself." Definitions of usability range from the high-level conceptualization incorporated in the ISO 9241 standard (Karat 1997) to more focused descriptions that include notions of user relevance, efficiency, user attitude, learnability, and safety (Lecerof and Paterno 1998). In detailing their concept of usability, (Lecerof and Paterno 1998) underscore that the most critical aspect of usability is contingent upon the actual system. For example, ease of use might be a primary criterion for systems designed for use by children, while efficiency is likely to be a major usability goal in the design of banking systems".

Usability can be defined as the measure of the quality of a user's experience with a product or system. This is the definition favored by the U.S. Department of Health and Human Services because of its general nature. Usability has also been defined as being "about learnability, efficiency, memorability, errors and satisfaction". The ISO-9241-11 standard defines usability as the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use. The standard further defines the terms used in this definition as follows:

- Effectiveness: the accuracy and completeness with which users achieve specified goals.
- Efficiency: the resources expended in relation to the accuracy and completeness with which users achieve goals.
- Satisfaction: freedom from discomfort and positive attitude to the use of the product.

- Context of use: characteristics of the users, tasks, and the organizational and physical environments.
- Goal: intended outcome.
- Task: activities required to achieve a goal.

The rationale behind this broad definition is that usability is about supporting users in achieving their goals in their work, and is not only a characteristic of a user interface.

More specifically for information systems such as e-commerce applications, usability is generally regarded as ensuring that interactive products are easy to learn, effective to use, and enjoyable from the user's perspective, and involves the optimization of user interaction (Lecerof and Paterno 1998).

From the user's perspective, perceived ease of use is the degree to which a person believes that using a particular system will be free from effort (Davis 1989).

From a systems theory perspective, since every system action uses resources, usability is the minimizing of the resource costs. In cognitive systems, the most significant resource is human effort (Whitworth and Zaic 2003).

4.8.2.2 Description. Usability is related to the tasks that need to be accomplished by a user. It therefore depends on the context of both the task and the user. Thus the characteristics that make a system usable may differ with different sets of users. Moreover, those characteristics of a system that make it usable for one group of users may cause it to be unusable for another group of users (Goodwin 1987).

Usability principles are usually fairly broad and general. This makes it difficult for non-usability experts to apply these principles to a specific domain (Singh and Kotzé

2002). Thus, in order to effectively incorporate them, it is necessary to understand and specify the context of use of the information system. This is defined by the characteristics of the users, tasks and the organizational and physical environment of the system (Jokela, Iivari et al. 2003).

4.8.3 Conclusions

Usability is probably one of the most well documented software requirements. It has however had very diverse definitions. This might have to do in part with the fact that usability depends on the user, the task, the application, and the environment in which the user employs a given application to perform a particular task. The layman's definition of usability obtained from a Standard English dictionary is more concerned with usability as a measure of convenience of using an application, while the definition adopted by the WOSP model is more concerned with the usability of an application in the context of minimizing the resources used for a given task.

4.9 Review Conclusions

This literature review begins by taking stock of arguments for the systematic evaluation of software by individuals in order to assess the acceptability of the application. To facilitate this assessment, it is critical that the software be comprehensively described in terms of its intrinsic qualities that would be of interest to the user. A survey of the more prominent models of user acceptance indicates that intrinsic software characteristics taken together comprise a technological aspect, alongside others, that influences user acceptance of software.

Perceived usefulness and perceived ease of use proposed by the TAM are two technological characteristics of software that have been found to influence the user acceptance of information systems. This survey of IS literature finds that there exists more than just these two factors characterizing the intrinsic qualities of information system software. Also, several other models have been proposed that combine several of these factors in various ways in order to describe the overall qualities of information systems.

The review makes a comparison between the major existing models that explain the performance of information systems, and its impact on users, based on the intrinsic qualities of software. Of particular interest in this regard are the TAM and the WOSP model.

TAM is considered to be a performance evaluation model. In order to assess the performance of an information system of interest, the user gauges the extent to which he perceives usefulness and ease-of-use to be present in the system. For either of these factors to be perceived, they must be present in the system. In other words, they must have been designed into the system and be present. Something that is not present cannot be perceived. On the other hand, a factor that is present in the system but cannot be perceived by the user is not appreciated by the user. So the factor must be present and perceptible to be valued. In the case of TAM, the perceived usefulness and ease-of-use correspond to the functionality and usability designed into the system

The TAM was conceived in what was probably mainly a human-computer interaction context. In the years since then, information technology has changed considerably, with information systems becoming much more complex. A lot of this complexity has to do with the social setting in which information systems are increasingly used. For example, in addition to the human-computer interaction, information systems are increasingly used to convey information and meaning between individuals and in groups of individuals that may have an identity, such as different organizations or online communities. This has led to the information systems being thought of more as socio-technical systems.

The WOSP model is derived from systems principles (Bertalanffy, 1968). The model proposes eight factors that, in total, describe the performance of a system. It proposes that the performance of the system reflects the extent to which these factors are present. Thus if these factors are present in a system, this should be apparent to the end user. Therefore, while the WOSP model may be viewed as a design model, the end user should be able to perceive those factors that are designed into the system. It then follows that the user should be able to use these factors as criteria in evaluating the performance of a system as to the extent to which he perceives the factors of interest to him as being present.

The WOSP model and TAM can be linked via the notion of system levels as applied to computer systems. (Whitworth and Zaic 2003) propose 4 levels:

- Mechanical: comprising the computer hardware system, including printer, keyboard, mouse, etc.
- Information: This is the software system, and comprises the mechanical system in combination with software.
- Cognitive: This includes personal attitudes, beliefs, opinions and ideas.
- Social: This involves group norms and culture.

These levels are given with increasing level of complexity, with the mechanical level being the simplest, and the social level being the most complex. Also, each successively higher level includes the lower level.

Hence, while the TAM can be seen as dealing with information systems at the cognitive level, the WOSP model describes information systems at the highest level as socio-technical systems. Also, the WOSP model not only includes system criteria relevant under TAM, but adds criteria corresponding to the complexities of socio-technical systems as well. This leads to the first conclusion from the literature review.

Conclusion 1: The recently proposed WOSP model, which is based on systems theory, is more comprehensive than the other existing models, in that it incorporates more intrinsic qualities of software. Furthermore the WOSP model encompasses the other existing models, so that these others can be subsumed into the WOSP model.

Using the eight intrinsic software qualities or factors proposed by WOSP as a framework, this review does a comprehensive survey of the literature and organizes the different definitions of the system qualities according to those WOSP factors. And in doing so, the review has confirmed that indeed the intrinsic software qualities proposed in the WOSP model are well documented in IS literature. A notable exclusion from the factors is a construct that explicitly deals with the intrinsic cost or economic value of the software. Although cost is mentioned in the literature, there is no clear explicit construct. A second major conclusion from the review is as follows.

Conclusion 2: The eight WOSP software characteristics are confirmed as valid and well-documented intrinsic software features. Each of them is however defined in a variety of ways, sometimes with little common agreement.

The third conclusion has to do with the presence of a number of qualities in software.

Conclusion 3: Given that each of these criteria could possibly influence a user, which has the greater influence likely depends on the specific application and the use to which it is being put. Moreover, since the qualities interact and modify the effect of each other, it is likely that the effect of any given quality in an application is actually not as high as it would be if it was the only criterion under consideration.

Of research interest is how users evaluate the performance of information systems as a prelude to acceptance of the software. TAM is the only model so far that has investigated how software qualities affect user's choice by investigating the influence of perceived ease of use and perceived usefulness of software on user's acceptance of the software. My research seeks to expand on this. Thus, the research issue is whether users take account of the eight aspects of intrinsic software quality proposed by WOSP, namely security, extendibility, flexibility, reliability, functionality, usability, privacy, and connectivity when they are assessing the performance of alternative information system software for purposes of acquisition, adoption, or just choosing between alternatives.

4.10 Research Question and Hypothesis

This section looks at the research question that arises from the literature review, and subsequently proposes the hypotheses that will be experimentally tested.

4.10.1 Research Question

For various reasons today, users of information systems are called upon to evaluate the performance of applications. This could be for purposes of acquiring software, or for purposes of selecting between available alternatives such as browsers. One of the key issues influencing the evaluation process of users is the intrinsic technical quality of the application under consideration. The WOSP model proposes that this technical quality is composed of eight factors, namely, security, extendibility, reliability, flexibility, functionality, usability, privacy, and connectivity. Of note is the fact that connectivity and extendibility are 'social' factors. Bearing this in mind, a pertinent research question arising from the foregoing review would be:

In the context where an individual is evaluating one or more information systems for use by an organization, how well do the eight WOSP criteria apply?

In order to answer this question, it must first be demonstrated that users indeed understand each of the WOSP constructs. This is especially important given the proliferation of definitions existing for each of the criteria. Having done this, it is important to show that users actually prefer to be able to consider any of the eight criteria when evaluating software. And finally, if they prefer to be able to use all eight, it is important to demonstrate that they consider at least several of them to be of equal significance for the evaluation of software. Taking the above into consideration, the research question can further be detailed into the following questions:

- Do users understand each of the WOSP criterion goals as aspects of system performance?
- Do users prefer to select from the set of all eight WOSP criterion goals when evaluating software in general? i.e. is each criterion goal considered important in general software selection?
- Do users rank the criteria differently for a given type of software?

4.10.2 Hypotheses

The constructs that represent the eight evaluation criteria suggested by the WOSP model are well documented in the IS literature. However, for them to be helpful to users for the evaluation of given software, it must be possible to have a clear operationalization of the constructs. Furthermore, it must be possible to operationalize the construct in a manner that ensures content validity so that users can see the correspondence between the operationalization and the definition of the constructs. Finally, users must find that the operationalization is an important measure of the constructs. This gives rise to the following hypotheses: For each given criterion,

- H1a: Users in general will understand the operationalization
- H1b: Users will in general consider the operationalization a valid measure
- H1c: Users will in general consider the operationalization to be important

The WOSP model suggests a number of software evaluation criteria, such as privacy and security in addition to functionality and usefulness, which are by far the most common evaluation criteria. The two correspond to the very commonly encountered and applied constructs of usefulness and ease of use, respectively, from the TAM model. In addition to the above hypotheses, if the criteria expansion by WOSP is valid, subjects should prefer to use it. They should also feel more confident using the WOSP criteria, believe the evaluation using the WOSP criteria was more accurate and complete, and be more satisfied with the outcome. Moreover, if the other criteria incorporated in WOSP are seen to be as important for software evaluation as functionality and usability, users should prefer to use them at least as frequently as they would prefer to use functionality and usability for the evaluation of software. This gives rise to the following hypotheses:

- H2:Users will prefer to use the WOSP criteria rather than just functionality and usability to evaluate software for an organization, and more specifically feel:
    - H2a:More confident in their choice using WOSP than just functionality and usability
    - H2b:That the WOSP evaluation was more accurate than just functionality and usability
    - H2c:That the WOSP evaluation was more complete than just functionality and usability
- H2d:Other WOSP evaluation criteria will be at least as frequently used as functionality and usability are on average

Furthermore, if the WOSP model in fact improves on usability and usefulness for software evaluation, some of the other factors should play a significant role in technology acceptance. If they rate the same as, or higher than, the WOSP functionality and usability criteria which correspond to the very commonly encountered and applied constructs of usefulness and ease of use from TAM, then the WOSP additional criteria are useful. However, if the top evaluation criteria are functionality and usability, and other factors are much lower, then the WOSP model's contribution is minimal. This suggests the following hypothesis:

- H3a: in the WOSP evaluation, performance criteria other than functionality and usability will contribute significantly, and so be seen as important, to technology acceptance.

Ideally, if the performance factors are equally important to users, all the eight performance factors should be perceived to have an equivalent effect on the performance of software. So if the total performance of software is taken as 100%, they should each account for 12.5% of the evaluation decision, since there are eight of them. However, the WOSP model predicts that the extent to which the individual performance factors are important depends on both the system and the environment in which it is operating, meaning that the factors are all not necessarily of equal importance. This leads to the following hypothesis:

- H3b: All eight factors are not considered to be equally important by users.

4.10.2.1 Operationalization Check Hypotheses. Tying this last set of hypotheses with the first set, in the event that the clarity, validity, and importance of the operationalizations of the eight criteria is significantly different, it is logical to suppose that this would have an effect on the importance that users attach to each of the criteria. This leads to the following set of hypotheses:

- H4a: Criteria operationalizations that are clearer will lead to higher importance being attached to the respective criteria.

- H4b: Criteria operationalizations that are more valid will lead to higher importance being attached to the respective criteria.

- H4c: Criteria operationalizations that are more important will lead to higher importance being attached to the respective criteria.

CHAPTER 5

METHOD

An exploratory experiment was carried out to investigate the main research question just described at the conclusion of the previous chapter: In the context where an individual is evaluating one or more information systems for use by an organization, how well do the eight WOSP criteria apply?

5.1 Pilot Experiment

5.1.1 Approach

5.1.1.1 Research Strategy. The research strategy involved the two issues below:

- Obtaining software with sufficient variance in the WOSP parameters.
- Getting subjects to rate the software on the basis of the WOSP parameters.

A major hurdle in the experimental design was obtaining enough variance in the WOSP parameters. If the subjects had been presented with real software, the software would likely be adaptive and successful; otherwise it would not be there in the first place. However, all such software would have similar, and/or generally high performance, and would therefore not have sufficient variance. Moreover, it would take the subjects some time to evaluate and rate the software, especially since some aspects such as reliability only become evident with time. It was therefore decided to leave the actual software to the imagination of the subjects, and present them with a situation where other people (users) were testing each software and making a report. The next issue was bias. If the subjects were told the rating of each WOSP dimension for the various software, the purpose of the investigation would be obvious, and it might be argued that the subjects simply went along with the ratings given. Thus, it was decided to embed the rating of the various parameters in what would appear to be a normal text report. The subjects would then read the report on the software, and on the basis of the qualitative assessment of the users, rate the software. The reports were generated by computer, with the comments on each of the dimensions being randomly varied. A particular problem resulting from this was that the same phrases appeared several times over, which might have reduced the realism of the experiment.

5.1.1.2 Variables. Below are variables in this experiment.

Independent variables: There were several independent variables in this repeated measures design:

- How the subjects perceived each of the eight WOSP parameters: functionality, usability, security, extendibility, reliability, flexibility, connectivity, and privacy.
- The types of software under evaluation. There were three: bulletin board, email system, and web browser.
- Since each subject evaluated each of the three software types, the order in which the three types were evaluated constituted the third variable.

Dependent Variables: The dependent variable was how subjects rated the software. This was in the form of recommendation of various strengths on the following five-point scale.

R: Recommended – we should buy it

R?: Recommended (with a few reservations)

RR: Recommended with some reservations

NR?: Not recommended probably, e.g. unless no better option ("soft" reject)

NR: Not recommended clearly ("hard" reject)

Control Variable: A ninth software aspect, alongside the eight WOSP parameters, was introduced as the control variable. This was whether the software was written in Java. In theory, this should have no effect on the final rating. Thus if, say, security predicted recommendation, but use of java did not, this would imply that security is a factor in the evaluation of software.

5.1.1.3 Subjects. The subjects were 43 graduate students at New Jersey Institute of Technology (NJIT) taking the course on Evaluation of Information Systems. Their average age was about 27 years. There were 18 female and 25 male participants. They came from a variety of cultural backgrounds. They opted to participate in the experiment as an option for the partial fulfillment of an assignment.

5.1.1.4 Task. Each participant was asked to evaluate reports on three alternative products having different values of the WOSP parameters for each of the three software types. For each product, the participant was to assign a rating, and then rank the three products for each software type.

5.1.1.5 Experimental Procedure. The following is a brief description of the procedure carried out for the experiment.

1. Students were asked to indicate their willingness to participate in the experiment by emailing the experimenter.
2. The participants were then emailed a brief description of the experiment and their role, along with a consent form for them to fill in.
3. As their first task, participants were asked to fill in questionnaire on their personal profile, which included aspects of attitude thought to be relevant to the experiment.
4. The participants were asked to assume the role of a manager in a firm and to recommend for purchase software on the basis of reports by six members of staff (A –F) working under him or her, who had had the opportunity to evaluate the software.

5. There were three types of software: bulletin board, email system, and web browser. For each, the manager received a report for three alternative products in the form of comments by all six members of staff. A typical comment by a staff member for a given product is as follows: *"One thing about this software is that it can easily add in a variety of third party plug in tools. Seeking a wide variety of opinions on this software would be useful. I found I could use it easily, with very little effort. I didn't assess how easy it was to customize the software. It was hard for me check how useful the software is, as I don't really know how to do this. From what I could see, this software can resist unauthorized tampering by some outside person. The software let me hide/restrict private information that I did not want others to see. I am fairly sure the program was written in java. I didn't test the software's network ability, as I do not know enough about it."*

Each of the eight WOSP dimensions was measured on a scale of 1-6, depending on how many of the six company reviewers made supportive statements on that dimension. For each of the dimensions, the user could either make a positive statement or a statement that essentially gave a blank. An example of a positive statement for the security dimension would be "If there was a security breach in the software I would be notified." A blank statement would be "With regards to the system's software, it would be better for you to see what other people say." The presence of such blank statements was explained to the participants as being due to the fact that the company reviewers were not expert in all areas. It was expected that when all six of the reviewers made positive statements about the dimension, participants would conclude that the software product was very secure, and the fewer reviewers made such statements, then the less secure the product was. Then if security influenced the recommendation by the subjects (managers), there would be a relation.

In addition, subjects were required to justify their choices, giving qualitative data. Finally, they were asked to rank the 3 products of each software type.

5.1.2 Analysis and Results

There were 43 participants in this exploratory experiment. Each participant evaluated three alternative products for each of the three software types. Thus each participant evaluated nine different types of software, giving a total of 387 different software evaluated.

The primary statistic of interest in this experiment was the correlations of the WOSP parameters with software rating. Table 5.1 below gives a summary of the correlation between the various WOSP parameters and the overall rating of software. It also gives the P-values of the correlations

Table 5.1 Correlations and P-Values of the WOSP Parameters

	Functionality	Usability	Security	Extendibility	Reliability	Flexibility	Connectivity	Privacy	Control
R-value	0.011	0.098	0.172	0.118	0.182	0.067	0.125	0.064	-0.039
P-value	0.836	0.053	0.007	0.02	0.0003	0.192	0.014	0.205	0.444

From the results, it can be seen that security, extendibility, reliability, and connectivity are correlated to the overall software ranking to a significant level.

5.1.3 Limitations of Experiment

5.1.3.1 Interpretation of Performance Parameters. It appears that software properties associated with some parameters were sometimes taken as being characteristics of other parameters. For example, positive statements made in regard to privacy were taken to be statements about security. Similarly, statements on connectivity were attributed to extendibility, and vice versa. This could also have been the case of flexibility being perceived as an aspect of usability.

An interesting observation with regards to functionality specifically was the fact that there was no correlation between the parameter and the overall software rating. This was similarly the case for usability, though to a much lesser extent. A possible explanation for this could have been the inability of the subjects to perceive these two parameters, as presented, as distinct properties separate from the other WOSP properties. Thus, if the software was perceived to have properties that were considered desirable, it was perceived to be functional, and conversely, whether or not it actually had a high level of functionality. And the same applies to usability. The properties of functionality and usability were possibly attributed to other parameters.

5.1.3.2 Decision-making Process. A major assumption of the experiment was that subjects would evaluate the software by cumulatively taking account of the presence of the various WOSP parameters. This in fact did happen is some fashion or other in many cases. As an illustration of this, one subject explained his rating procedure as follows: "As with the previous two software categories, I arranged the substantive comments of the evaluators in spreadsheets. Product 1 had 16 distinct comments for a total of 19 substantive comments. More than one evaluator may have had the same comment, hence more total comments than distinct comments..." Another subject said: "This product got the best overall reviews in the areas of compatibility, flexibility, security, ability to maintain privacy, functionality, stability and self maintainability. More information on usability would further promote the product's case. One of the strong points is that this software is extensible because the architecture conforms to open source standards."

However, there was a wide variation in how the subjects went about assessing the software. Some examples illustrating this diversity are given below.

In some cases, the subjects seemed concerned with only certain performance parameters, which were then their main criteria for selecting and ranking the software. As an example of this, one subject stated the following: *"The only good thing visible about this product was its flexibility to adapt to outside changes and its user friendliness. The software was also compatible with other system devices and peripherals. The scalability of the software was in question. The most important factors that should be present in any Email software are privacy and security..."* Another had this to say: *"I give this software my highest recommendation. It scored very high with all of the employees who tested it. They seemed enthused with this product because they seemed to supply the most feedback. Many of the respondents commented that this software provided a lot of privacy. In fact eight privacy issues were mentioned in the summaries. Three comments were made on security, one employee said it did not allow hacking, another commented on its login page and the third mentioned that it protects data. As a manager, security is one of my top priorities. I thought that this product provided the adequate protection the organization needs but because I realize that these employees are not experts in the field, I still have my reservations. I wonder if they are truly qualified to access security issues..."*

Another subject said this: "There were two primary reasons for ranking this software as a number one buy. It's open source, which will be important if we decide to customize the software. The other reason was the software's ability to do a maintenance check to fix file and/or data problems. Most of the other important features, such as security, privacy, efficiency, interoperability, preferences menu, productivity, and adaptability seem to be present in all three products."

5.1.3.3 Experimental Method. Some subjects were more concerned with, and distracted by, the quality of the reviews by the users, with some getting upset at the perceived lack of information. One said: *"No one seems to know what the hell is going on with this software. Either they are a self-proclaimed "expert" or just don't know how to answer the question. In this case I am throwing the baby out with the bath water and not taking anybody's word."* and *"Am I in bizarro world?? These so called recommendations sound like the first sorry lot. I have a feeling that this company is in a heap of trouble using this methodology. There was nothing here that would make me want to sign a contract for this software."* Another said: *"66% of the respondents did not give the system a thorough investigation. There is no mention of security. Unless we get a more thorough testing, I will pro-actively work to strip this system of its candidacy."*

5.1.4 Suggestions for Future Improvement of Experiment

1. There seemed to be some confusion as to the meaning of some of the concepts among some subjects. This suggested that the WOSP concepts needed to be better operationalized so that the intended meaning by the researchers was actually what the subjects understood.

2. The operationalization depended to some extent on the type of application being considered. This is particularly the case with functionality (effectiveness), which is largely application-specific.

3. Given the fact that some subjects seemed to attach greater value to some performance parameters than to others, it might be insightful to have the subjects give an indication of the importance they attach to the various parameters, and whether indeed all the parameters were important for the selection of software.

4. An important first step in improving the experiment would be to gain a better, more systematic understanding of the decision-making process of subjects especially as it relates to the evaluation of software. This would then be factored into the design of the experiment.

5. The statements by the reviewers, particularly the blank statements could be modified to be more realistic. This might include the use of negative statements

with regards to the various performance parameters. As performed in the experiment, the blank statements seemed to create quite a distraction in some cases, with the subjects devoting attention to this fact, to the detriment of the overall quality of their response.

6. An option to the above step might be to have subjects do the actual rating of software themselves, rather than rate on the basis of reports from reviewers.

7. Although there was nothing fundamentally wrong with the experimental method, it was empirical, and not a tried and tested technique. It might therefore have undetected and unacknowledged weaknesses. It would be very much preferable to use an established method whose usage and interpretation is well understood. This might actually be the most critical issue in the whole research. For example, how to combine the various predictor variables presented a particular challenge. Also to be taken into account is the fact that the predictor variables in reality are not just either present or absent, but rather vary in the amount to which they are present.

5.2 Experimental Subjects

Subjects will be students at NJIT taking a graduate course in the area of information systems. Participation in the experiment will be offered as an option to an equivalent course assignment.

5.3 Experimental Task

For research question 1 and research question 2, the subjects will be required to take on the role of a middle manager in an International company that has to select software for the entire firm. In this role, they will be requested to evaluate the operationalizations of the WOSP criteria for research question 1, and to respond to a questionnaire for research question 2.

For research question 3, the subjects will be required to take on the role of senior IT purchasing manager in the same company. In this role, they will be required to evaluate many alternatives of software for possible purchase using the eight WOSP criteria.

5.4 Research Question 1

Do users recognize and understand the WOSP criteria? Arising from the analysis of the exploratory experiment, it is of critical importance to investigate whether users actually recognize and understand each of the WOSP criteria as distinct concepts. The methodology described below is adopted for research question 1.

5.4.1 Methodology

5.4.1.1 Experimental instrument. The development of the instrument is described briefly below

i. For each WOSP factor, the IS literature was reviewed for approximately fifteen general descriptive statements of the factor that could be used to operationalize the criterion. This ensured content validity, an issue that was apparent from the exploratory experiment.

ii. The exploratory experiment had suggested that some of the manifestations of the WOSP criteria were application specific. It was therefore decided to use just one type of experimental software. These descriptive statements were adapted to suit the illustrative application to be used in the research, in this case the browser. The web browser was chosen as the illustrative software because of its increased importance as a sophisticated socio-technical software application in everyday and pervasive usage in many different ways. This variety of usage would in all likelihood be impacted by the WOSP factors if indeed these factors are of consequence. It is also likely that all prospective subjects for the experiment would be at least fairly familiar with browsers.

iii. For each factor, the definition, synonyms and the statements were listed together on one page. Five individuals with considerable and varied experience with IT in general and browsers in particular were asked to rate each statement on a scale of 1-5 for clarity of meaning; validity as an illustration of the factor, given the definition; and importance of the statement for assessing applications. In addition, they were asked to give any other helpful comment they might have. The individuals included 2 reference librarians, each with over 10 years experience using the browser as a major tool in their everyday office work. The remaining three were PhD students with considerable knowledge of IT in general and browsers in particular. On the basis of this feedback, the statements were amended together with the presentation.

iv. The resulting statements were further discussed with one of the initial group of experts who, apart from considerable computer science experience was exceptionally knowledgeable about many technical and user aspects of browsers. Again, the statements and presentation was amended. The resulting statements are given in Appendix A.

5.4.1.2 Variables. The independent variables will be the various statements for each the criteria, while the dependent variables will be the rating of each of these by the subjects.

5.4.1.3 Procedure. Subjects will be presented the statements for all the criteria via email and asked to rate the statements on a scale of 1-5 for each of the following

- Clarity of the meaning of the statement.
- Validity of the statement relative to the given definition of the factor.
- Importance of the statement for assessing software.

The order in which the subjects are presented the statements for each factor will be randomized using the random function in the EXCEL spreadsheet, so that no two subjects receive the statements in the same order. This will serve to control for the effect of order.

5.4.2 Analysis

After the subjects have rated the statements for each of the factors along the three dimensions listed above, the score for each statement by each subject for each dimension will be added up. Thus each statement of a given factor will have a total score for the ratings by all subjects for clarity, validity, and importance.

For each factor, the average score for all subjects for each of the variables clarity, validity and importance will be calculated. Further analysis for each will be as follows:

5.4.2.1 Reliability. Cronbach's Alpha reliability test will be done on the scores of clarity, validity and importance. This will give a measure of the correlation between the statements, and to what extent they apply to the same construct.

5.4.2.2 Clarity, Validity and Importance of Criteria. A single sample t-test will be performed between the average score and the midpoint of the scale. This will indicate whether the subjects recognize and understand the particular factor.

For each factor, the statements will then be sorted in descending order of the average individual score, first by importance, then by clarity, then by validity. It is assumed that those statements that ranked highest are those that the subjects are most in agreement with. For the purposes of this research, the six statements that rank highest for each factor will be taken as the most descriptive for the factor. They will be used for priming in the investigation of research question 3.

5.5 Research Question 2

Do users prefer to select from the set of all eight WOSP criterion goals when evaluating software in general? In other words, is each criterion goal considered important in general software selection? An analysis of the self reports of the exploratory experiment indicated that users might vary in the number of criteria they choose to use in the selection of software. This issue is addressed more specifically with this research question.

5.5.1 Methodology

The two most common and well known software performance factors are usability and functionality. The use of only these two to evaluate software will be compared to using all WOSP evaluation criteria.

5.5.1.1 Questionnaire Measures. These are described briefly below. The detailed instruments for this research question are given in Appendix B.

a) Method preference will be measured by presenting the question: "In general, if you were evaluating software for a company, would you prefer to choose the criteria to evaluate with from a) all eight WOSP criteria, OR would you prefer to regularly choose from b) a set of fewer criteria?"

b) In addition, the following attitude question items will be measured on 5-point Likert scales: "Suppose you only used the criteria Functionality and Usability to select software. How would this compare in the following ways with using the eight criteria WOSP to select software?"

   1. Confidence: How confident are you that the method's data will give the best software choice? (Very Confident to Very Unsure)

   2. Accurate evaluation: How accurate is the method in testing software performance? (Very Accurate to Very Inaccurate)

   3. Complete evaluation: The method includes all the dimensions I need in order to evaluate software. (Strongly Agree to Strongly Disagree)

c) Finally, the following question on the frequency of use will be asked: "Supposing you were evaluating many different types of software. For each of the following criteria, state how often you would expect it to be a critical evaluation factor."

5.5.1.2 Variables. The independent variables will be the various questions in the questionnaire, while the dependent variables will be the subjects' responses to these.

5.5.1.3 Procedure. Subjects will electronically be sent the instrument shown in Appendix B, together with detailed instructions on how to fill in the questions in the instrument for collecting the data described in the previous section.

5.5.2 Analysis

T-tests will be calculated to compare the WOSP criteria with the commonly encountered functionality and usability duo.

5.6 Research Question 3

How do the users rate these criteria relative to each other for purposes of system evaluation?

5.6.1 Choice of Methods

The research involves the study of how information system performance is perceived to vary with the eight WOSP variables, or how each of the eight criteria influences the rating of software from the perspective of a user. The research problem would therefore involve some form of multivariate analysis. According to available literature (Hair et al., 1995), the selection of the appropriate multivariate technique to use for analysis depends on three main guidelines.

1. Whether the variables can be classified into dependent and independent variables.
2. If so, how many variables can be treated as dependent in a single analysis.
3. How the variables are measured.

5.6.1.1 Classification of Variables. If the variables can be classified into dependent and independent, then a dependence technique should be used. This is a technique whereby the dependent variable or variables is, or are, predicted by the independent variables. If no such classification of variables is possible, then an interdependence technique should be used, whereby all the variables are analyzed simultaneously.

In the case of the experiment, the variables could easily be categorized into the two classes, with the measure of user evaluation of the performance of the software constituting the dependent variable, while the eight WOSP parameters of connectivity, functionality, flexibility, extendibility, privacy, reliability, security, and usability constituting independent variables. The most appropriate technique is therefore a dependence multivariate technique.

5.6.1.2 Number of Dependent Variables and their Measurement. Multivariate dependence techniques can further be classified on the basis of the number of dependent variables. In addition, the techniques can be classified as those with either metric (quantitative or numerical) or non-metric (qualitative/categorical) dependent variables. If the research problem involves a single dependent variable that is metric, the appropriate analysis techniques would be either multiple correlation analysis or regression. If the single dependent variable is nonmetric (categorical), then multiple discriminant analysis or linear probability models would be more appropriate. If however the problem has several metric dependent variables, multivariate analysis of variance and canonical correlation analysis are appropriate. If the problem involves several nonmetric dependent variables, canonical analysis can be used after transforming the variables through dummy coding. Conjoint analysis technique is a special dependence technique that can be used for a single variable that is metric or nonmetric.

This research involves a single dependent variable that is metric. So multiple correlation analysis, regression, or conjoint analysis could potentially be used.

Below is a chart that summarizes the selection procedure of the various multivariate techniques from (Hair et al., 1995).

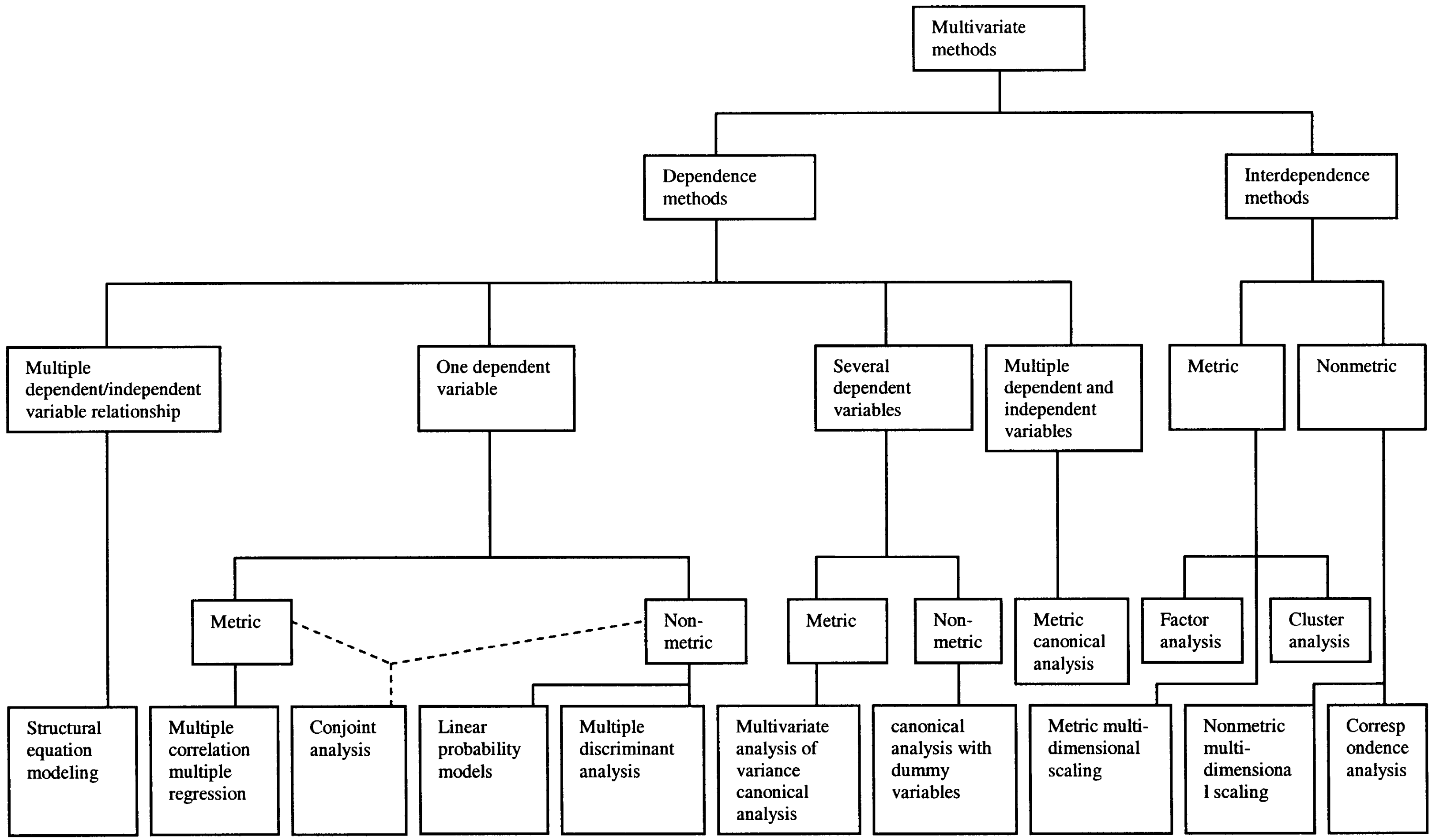

Figure 5.1 A roadmap of multivariate methods.
(Source: Hair et al, 1995)

Multiple Regression: While using regression analysis, a decision has to be made on the number of predictor variables to be included. The assumption is that each additional predictor variable gives more information and thus improves the prediction of the dependent variable.

Conjoint Analysis: This method of analysis arose out of a need to analyze the effects of predictor variables that are qualitatively expressed, or weakly measured. It portrays a user's decision realistically as a trade-off among several attributes or factors. Conjoint analysis is unique in that it is able to accommodate a metric or non-metric dependent variable, and that it is able to make use of categorical predictor variables. The major objectives of conjoint analysis include:

- Determining the contributions of predictor variables, and their respective values to the determination of consumer or user preferences.
- Establishing a valid model of consumer preferences that can be used to predict user acceptance of any combination of attributes, including those not originally evaluated by consumers.

5.6.1.3 Conjoint Analysis versus other Multivariate Techniques. Conjoint analysis has three distinct differences with other multivariate methods:

- It is decompositional in nature.
- Estimates can be made at the individual level.
- It is flexible in terms of the relationships between dependent and independent variables.

Compositional vs Decompositional Techniques: Using conjoint analysis, once a subject's overall preference of an object of given characteristics is known, this preference is decomposed to determine how much is due to each factor and each possible value of the factors. In contrast, with compositional techniques such as regression, ratings are collected from subjects on many characteristics of a model of interest. These ratings are then related to some overall preference rating, resulting in a predictive model. In so doing, the subject's overall preference is composed from the evaluations of the product on each factor.

Estimates at the Individual Level: Conjoint analysis is unique in that it allows for the generation of a preference model for each subject. This is in contrast to other multivariate techniques which develop a common model for all the respondents. In conjoint analysis, estimates can be at either the individual or the group level.

5.6.2 Methodology: Conjoint Analysis

A conjoint analysis experiment typically involves the following outlined steps (Hair et al. 1995).

1. Statement of the research issue
2. Design of stimuli
   a) Specification of factors and levels
   b) Specification of the basic model form
   c) Creation of stimuli
3. Data collection
   a) Selection of the presentation method
   b) Selection of a measure of preference
4. Estimation of the part worths
   c) Selection of the estimation technique
   d) Evaluation of the results
   e) Interpretation of the results

5. Application of the conjoint results

These steps are subsequently elaborated on below.

5.6.2.1 Description of Conjoint Analysis. Conjoint analysis enables an understanding of how subjects develop preferences for products or services. It is based on the idea that people evaluate the value of a product or service by adding up the separate amounts of utility provided by each of the attributes of the product or service.

A unique feature of conjoint analysis among multivariate analysis methods is the fact that a set of hypothetical products or services are first constructed by combining the attributes that make up the product or service at various levels. The hypothetical products and services are then presented to the subject who is then required to indicate his preference from among the set as he practically would in real life. Based on the subject's preferences, the technique is able to determine the importance to that individual of each attribute and attribute level.

For the technique to be applied successfully it must be possible to describe the product or service fully in terms of its attributes/factors and the various levels these attributes can attain. A product or service with a particular set of levels or values of the various factors is referred to as a treatment or a stimulus.

Analysis of a subject's evaluation or preference of specific combinations of treatments gives the subject's preference structure, which shows not only how important each factor is to the individual's decision, but the importance of the various levels within each factor. In ranking the combinations of treatments, the individuals must make a tradeoff as they must take the factors and levels they like as well as those factors and levels that they do not like in a particular stimulus and make an overall judgment. The overall preference for a particular stimulus can be regarded as the total worth of that product or service. Then the factors in their various levels would be considered part-worths of the product. This can be represented in the general form of a conjoint model as follows:

Total Worth for product = Part-worth of $level_i$ for $factor_1$ +
part-worth of $level_j$ for $factor_2$ + ...+
part-worth of $level_n$ for $factor_m$

Where the product or service has m factors, each with two or more levels. The particular stimuli consists of $level_i$ of $factor_1$, $level_j$ of $factor_2$, and so on up to $level_n$ for $factor_m$.

5.6.2.2 Research Problem. The research issue is how users of software rate the WOSP criteria, namely connectivity, functionality, flexibility, extendibility, privacy, reliability, security, and usability for purposes of system evaluation.

It is reasonable to assume that these factors comprehensively describe the performance of software, and are thus the main ones that the users need to take into account when evaluating the software in terms of its performance. So the factors are the key decision criteria in the evaluation and choice process for software. The emphasis is on understanding this decision process in terms of the importance to the users of each of the performance factors. In terms of the conjoint model, this is equivalent to comparing the part worths of the performance factors.

5.6.2.3 Designing Stimuli. A stimulus is a specific set of factors, each at a particular level, which is evaluated by subjects. Design of stimuli involves selection of the factors to include in the stimuli, and different levels that each factor can assume. It also includes selecting the actual number and composition of the different stimuli.

Selection of Factors: The factors to be used in the design of stimuli should, as far as possible, meeting the following assumptions and conditions.

1. All factors that potentially add or detract from the total worth of the performance of software should be included. This helps represent the user's judgment and evaluation process correctly.

2. Not only must the factors be important, but they also must have the quality of differentiating between the alternative objects of assessment, as these are the factors that are important in judgment decisions.

3. The factors included in the stimuli must be distinct and, in addition, represent a distinct concept. Also, the concepts should be precise and easily understood, so that there are no differences in the perceived meaning by different people; the only difference should be how the different individuals actually feel about the factors.

4. Multicollinearity among the factors should be avoided, otherwise there results in a lack of orthogonality among the attributes, whereby the part-worth estimates for the factors are not independent of each other.

5. The factors must be easily communicated so that the subjects can evaluate them realistically.

6. The number of levels for the various attributes should, as far as possible, be balanced. It has been observed that the relative importance of a variable increases with increase in number of levels.

According to the WOSP model, which is based on systems theory, the eight factors comprehensively define software's performance. Representing these factors in the stimuli, therefore, accounts for all the major factors that affect the total worth of the performance of software. This then presents a realistic situation for the subject's assessment. In addition, since it is these factors that make up the performance of software, it is the extent or level to which these factors are present that differentiates the various alternatives of a given application.

It is also critical to ensure content validity of the factors in question for this kind of research, and that the factors are precisely, easily and uniformly understood by all the subjects who evaluate the stimuli.

Selection of Levels: In order to avoid distorting the relative significance of the factors, they will all be given the same number of levels. Another consideration is the fact that there are eight factors in total, and so the levels that they could take need to be controlled so that the possible combinations for evaluation do not become too many. Moreover, the number of levels should not be to such granularity that the subjects cannot distinguish between one level and the next. Taking all this into account, three levels were selected for the factors: high, medium and low.

Selection of the Form of Model: There are two choices of conjoint analysis model for explaining a subject's preference structure based only on the overall evaluations of a set of stimuli. In the additive model, the individual simply adds up the part worths for each factor in a stimulus so as to get a total value for the stimulus being evaluated. The interactive model also assumes that the consumer sums the part worths to get the total of all the factors. In addition however, it allows for certain combinations of factors to be more or less than their separate combined worths.

The additive model is the most basic and common model, and accounts for 80-90% of the variation in preference in almost all cases, and is usually sufficient for most situations. It requires fewer evaluations from a user, and it is easier to obtain estimates for the part worths. It assumes only main effects for each factor with no interactions. Usually, the interactions predict much less than the additive effects, often not exceeding 10-15% (Hair et al. 1995).

An analysis of the feedback from the exploratory experiment suggested that subjects used the additive model in making their evaluations of a set of factors.

Based on this and the earlier comparison of the two alternative models, it was decided to use the additive conjoint analysis model.

Selection of the Part Worth Relationship: Here, the issue is how the different levels of a given factor relate to each other. In other words, what the relationship is of the part worths within a factor. In conjoint analysis, three alternatives are available:

- The linear relationship model, which is the simplest. It assumes that the part worths of the levels of a factor vary proportionately with the level.
- In the quadratic relationship model, the assumption of strict linearity is relaxed so that the part worths of the various levels of a factor have a simple curvilinear relationship.
- The part worth model, which is the most general.

For this research, the part worth model was chosen as it is the most general, allows for each level to have its own part worth estimate, and gives the most information regarding how an individual's preference varies with the level of a given factor.

Creation of Stimuli: Once the factors and the levels for the factors have been selected, the next stage is the creation of the stimuli for evaluation by subjects. Two main alternatives exist:

- The factorial design is used in situations where the number of factors and levels are sufficiently small. The total number of all possible combinations of levels of factors is correspondingly sufficiently small to allow the subject to evaluate each possible stimulus.
- The fractional factorial design is used in the alternative case where the number of factors and levels increases to the point that it is impractical for a subject to evaluate all the possible combinations of factors and levels and give consistent answers that are meaningful.

In the case of the fractional factorial design, the critical issue is to get a representative subset of all the possible combinations of factors and levels. For a factorial design, the stimuli must be an orthogonal subset from all the possible combinations that allows for the estimation of all the part worths of all the main effects. Interactions whereby the part worth for the level of a given factor is dependent on the level of another factor are assumed to be negligible. In an orthogonal subset, every level of a factor occurs with each level of another factor with equal or at least proportional frequencies so as to assure the independence of the main effects. It is the most parsimonious way of estimating the main effects, and ensures that information is not really lost by omitting some combinations. The conjoint module of the SPSS statistical software package was used to create an orthogonal fractional factorial design.

5.6.2.4 Data Collection. The objective of the data collection phase is to convey to the subject the various stimuli as realistically and accurately as possible and to obtain the subject's reaction.

Method of Presentation of Stimulus: The two main methods used for the presentation of stimulus in conjoint analysis are the trade-off method, and the full profile method. In the full profile method, each stimulus in a set is presented separately, and each factor in the stimulus has a level defined. It has the following main advantages:

- Each stimulus has all the factors that a user would have to consider, thus the decision-making process of a user involving tradeoffs among the various factors while making a preference is more realistic.
- There are fewer judgments to be made compared to the trade-off method, since each stimulus contains all the factors.
- Fractional factorial designs can be used with this method.
- The preference can be judged as either a rating or a ranking.

The following are some of the limitations of the full profile method:

- The judgment required of subjects is more complicated than with the trade-off method.
- As the factors increase, the possibility of information overload increases as the subject has to consider correspondingly more tradeoffs.
- The order in which the factors are presented to the subject might have an effect on the evaluation.

The full-profile method is recommended when the number of factors involved is six or less, whereas the experiment involved eight factors. However, the advantage of this method being more realistic and reflective of what decision a subject would be called upon to make regarding the evaluation of the performance of software outweighs the disadvantages. Also, the trade-off method would involve many more decisions that would most likely result in more fatigue to the subject than with the full-profile method. Taking this into consideration, the full-profile method of presentation will be used for the experiment.

For eight factors, each with three levels, the conjoint module of the SPSS software package produced a set of 27. In addition, six more stimuli were generated separately by the module to be used as holdouts so as to check for the consistency of the subjects' evaluation. Thus, the subjects will be required to evaluate a total of 33 stimuli. The set of stimuli is given in Appendix C.

In order to control for the effects of order, the order in which the stimuli in a set are presented to the various subjects will be randomized using the random function of the EXCEL spreadsheet software. Thus, no two subjects will receive the stimuli in the same order. Moreover, for each subject, the order in which the factors appear in the stimuli set will also be randomized using the EXCEL spreadsheet function. Again, no two subjects will have the performance factors in the stimuli set arranged in the same order.

Selection of a Measure of Preference: Using the full profile method of presentation, subjects can show their preference either by rating or by rank ordering. Each of these two measures of preference has advantages and disadvantages.

Rating on a metric scale has the advantage of being easy to administer and analyze. Also, it allows conjoint estimation to be done by multivariate regression. However, a major disadvantage is that subjects can be less refined in their judgment than if they were to rank the stimuli.

Ranking has the major advantage that it is likely to be more reliable, as the subjects are forced to be more refined and selective while sorting the stimuli. It has the disadvantage that it can get quite cumbersome the greater the number of stimuli.

For the experiment, a combination of both rating and ranking will be used as per the following steps:

- First, the subjects will be requested to grade each stimulus in one of five categories: strong, good, adequate, limited, and weak, based on the levels of the factors in the stimulus.

- Next, the worst, or least preferred, stimulus will be given a score of 1 and the best, or most preferred a score of 100. The rest of the browsers are to be given a score in between. However, the best stimulus in any of the categories above will be scored lower than the worst stimulus in the next higher category.

- Finally, stimuli will be ranked from 1 to 33, with 1 being the best preferred and 33 being the least preferred.

5.6.2.5 Variables. This section describes the various variables in the experiment.

Independent Variables: The following are the independent variables in the experiment:

1. The eight WOSP criteria of connectivity, functionality, flexibility, extendibility, privacy, reliability, security, and usability which made up the stimuli.
2. The order in which the 33 stimuli are evaluated.
3. The order in which the WOSP factors appear in each of the stimuli.

Dependent Variables: The dependent variable is how the subjects rank the stimuli from 1 to 33, or the preference of each browser as indicated by rank.

5.6.2.6 Procedure. The participants will be required to evaluate 33 stimuli comprising 27 experimental stimuli and 6 holdout stimuli. These stimuli represent different types and versions of web browser each with a unique set of the performance factors extendibility, security, connectivity, usability, privacy, flexibility, functionality, and reliability. The participants will rank all the stimuli according to preference. The subjects will be required to carry out the evaluation process as follows:

A Grade each browser Strong, Good, Adequate, Limited or Weak based on its ratings. There are 33 browsers, each with a separate ratings page. The browsers are not listed in any particular order.

B Score the browsers 1 – 100. Pick what is considered the worst browser and score it 1. Then pick the best browser and give it 100. Finally, give the other browsers scores in between 1 and 100. The browser grade may be used as a guide, e.g.:

i. Weak, score from 1- 20.
ii. Limited, score from 21- 40
iii. Adequate, score from 41- 60.
iv. Good, score from 61- 80
v. Strong, score from 81- 100

However the scores must be spread from 1 to 100.

C Rank the browsers from 1 to 33 in the Rank column, with 1 being the best and 33 the worst. No two browsers should have the same rank.

D Explain the ranking decision.

The entire experimental procedure will be carried out via email. From the pilot studies, it has been determined that both the instructions and the steps are sufficiently simple to be carried out by the subject. There is therefore no need for a face-to-face administration of the procedure. This has the advantage of standardizing the procedure and administration for all subjects, thus controlling for variability that might be introduced by the researcher. There is also the added advantage of simplifying the whole process, and also cutting down greatly on the time it would take to administer the experiment on an individual face-to-face basis, which in itself could introduce further variability. The instrument for this procedure is given in Appendix D.

5.6.3 Analysis and Interpretation of Results

Evaluating the Results: Conjoint analysis results need to be assessed for the accuracy of the estimated models at both the individual and the aggregate levels. The objective in assessing the reliability is to ascertain how consistently the model predicts across the set of preference evaluations given by each individual. For the rank-order data, correlations based on actual versus predicted ranking are used, such as Spearman's rho and Kendall's tau. These values are then tested for statistical significance.

As a further check, the accuracy of the model is measured against not only data from the original stimuli, but also with a set of hold-out stimuli. This is a set that is in every respect similar to the first or initial set. The hold-out stimuli are evaluated together with the other stimuli. However, data from the holdouts is not used for calculating the model parameters. Instead, the parameters estimated for the model from the experimental stimuli set are used to predict the preference for the hold-out stimuli. This predicted preference is then compared with the actual responses to assess model reliability. Specifically, this tests for internal consistency on the part of the subject, and shows whether the subject's decision model is consistently logical or not. A set of six stimuli will be used as hold-outs in the experiment.

In addition, with this kind of analysis, there is always the possibility that extreme values by any individual subject may bias the aggregate results. A boxplot analysis will therefore be done on the part worths of each of the eight factors generated by the subjects comprising the sample to check for any outliers. Of particular interest are extreme outliers, which are described as data values with virtually no chance of coming from the some population as the bulk of the data.

Assessing the Relative Importance of Attributes: Conjoint analysis gives an indication of the impact of each level of a factor by means of part worth estimates. It also assesses the relative importance of each factor. The part worth estimates are on a common scale, and so the greatest contribution to overall utility or preference, which is therefore the most important factor, is the one with the greatest range of part worths (that is, low to high). In order to provide a consistent basis for comparison across different individuals, the range values for each model are standardized by dividing each range value by the sum of all the ranges. Conjoint analysis calculates importance values for each factor such that the total for all factors comes to 100%. It is hence possible to compare the significance to users of the various factors.

Aggregate vs. disaggregate analysis and interpretation of results: The interpretation of the results of a conjoint analysis can be by either on individual subject results, which is referred to as disaggregate analysis, or on the aggregated results. In disaggregate analysis, the fit of each individual model is appraised relative to the assumptions of the model. An added advantage of this approach is that it enables the exclusion of data from subjects whose model parameters suggest that they did not perform the tasks properly (Hair et al. 1995). Both aggregate and disaggregate results are produced by the Conjoint Analysis software package.

In the aggregate analysis, one model is fitted to the aggregate results of all the responses. The two metrics by which the results of the experiment are given are adapted from (Bajaj 2000). In the first metric, the mean relative importance values of the criteria are presented. However, the mean is just a point estimate of the true value, whereas the confidence interval of the means is given as a better estimate of the true value of the parameter (Vining 1998). The 99% confidence interval will therefore be given for the average part worth.

The second metric takes account of the possibility that the average importance values could be biased by the extreme values of some subjects in the sample (Bajaj 2000). Thus, the percentage of subjects in the sample that had an average importance equal to or greater than 12.5% will be given for each of the criteria. Finally correlations will be calculated between the average importance of each of the criteria as obtained from conjoint analysis, and the average scores of clarity, validity, and importance for each of the criteria as obtained under research question 1.

CHAPTER 6

ANALYSIS

6.1 Subjects

Participation in the research was offered to graduate students in management as an alternative to regular course assignment. A detailed description of what it entailed was provided. The students who chose the experiment became the subjects from whom data was collected for analysis. They were however free to discontinue participation at any time that they wanted to do so.

There were 60 subjects involved in this study. They were graduate students from the School of Management at New Jersey Institute of Technology. They had, on average, been using browsers for over eight years, with the vast majority using browsers for at least five years. In the six months before the experiment, they, on average,e used browsers for 27.9 hours a week, with the vast majority using browsers for at least 15 hours. They reported to have been using browsers for various reasons, the most common being work-related, taking online courses, browsing for information, online purchases, and other financial transactions. The subject group was diverse in gender and culture.

6.2 Data Overview

This section describes how data collected was prepared for analysis. The reader may wish to skip this and proceed to the results in the next section.

6.2.1 Data Sets

There were three independent sets of data collected in this experiment, each corresponding to one of the three research questions as follows:

1. Recognition and understanding of the WOSP criteria. Specifically, of interest was whether the subjects found statements used for operationalizing the criteria to be clear, valid, and important.

2. Whether users prefer to be able to use any of all the eight WOSP criteria when evaluating software in general. Specifically, a) whether subjects would prefer to choose from all eight or from fewer, and if so, which ones; b) How subjects compared the eight criteria WOSP to the two common criteria of functionality and usability together for evaluating software, possibly for purposes of selection, in terms of confidence, accuracy, and completeness; c) How often subjects would expect each of the eight WOSP criteria to be a critical software evaluation factor.

3. How users rate the WOSP criteria relative to each other for purposes of information system evaluation. Specifically, how subjects rate the importance of the WOSP criteria relative to each other.

6.2.2 Data Collection and Preparation

This section briefly describes how data was obtained and prepared for analysis. In all cases, statistical analyses were carried out using the SPSS statistical analysis software package.

6.2.2.1 Recognition and Understanding of the WOSP Criteria. Subjects gave ratings on questionnaires for the clarity, validity, and importance of the statements operationalizing each WOSP criterion. These responses were then analyzed.

6.2.2.2 Preference for Using Any of All the Eight WOSP Criteria. a) Subjects were asked to indicate whether they would prefer to choose software evaluation criteria from either all eight WOSP factors or from a set of fewer criteria. They were further asked to state how many criteria, if fewer than eight, and to select which of the WOSP factors they would prefer.

b) Subjects were then asked to separately rate, on a Likert type scale of 1-5, using the 2-criteria combination of Functionality and Usability; and using the WOSP criteria to evaluate software on the following aspects:

i) Confidence that the method would give the best software choice.

ii) Accuracy that the method correctly measured software performance.

iii) Completeness, i.e. that the method covered all the relevant factors.

For each of the above, the responses by the subjects were then analyzed using SPSS.

c) Subjects were asked to rate each of the eight WOSP criteria on a Likert type scale of 1-5 as to the frequency with which they would expect it to be a critical evaluation factor. The ratings by the subjects for all the criteria were then analyzed.

6.2.2.3 Data on Users' Rating of the WOSP Criteria. Subjects were asked to rank 33 different alternatives of browser according to the subjects' preference. The rankings by each subject were then separately transferred from the questionnaire to an Excel spreadsheet. Here, the browsers, whose order had been randomized to minimize order effects, were reordered such that for all subjects the browsers were in identical order to facilitate further statistical analysis.

6.2.3 Missing Data

For all data sets, the raw data from the subjects was inspected, and if the questionnaire was not complete, the results were not used for further analysis. Therefore there were no cases of missing data in the analyses.

6.3 Recognition and Understanding of the WOSP Criteria

The analysis involved two steps. The first one was to check the reliability of the questionnaire instruments. This was done by calculating Cronbach Alpha values for the subject responses for all eight factors for each of the three aspects of clarity, validity and importance.

The next step was to perform a one-sample t-test to investigate the clarity, validity and importance of the questionnaire instruments for operationalizing the WOSP factors. Subjects had been asked to rate each of the statements for all eight factors on how clear, valid and important they felt the statement to be. The rating was on a scale of 1-5 as follows:

1=Strongly Disagree; 2=Disagree; 3=In the middle; 4=Agree; 5=Strongly Agree. The statistical analysis involved comparing the averaged response of the statements of an instrument with the rating scale midpoint of 3.

For each of the three aspects of clarity, validity and importance, the results of the statistical analyses are given below.

6.3.1 Clarity Aspect of the WOSP Criteria

The results of the instrument reliability and the one-sample t-test for responses regarding the clarity of the statement used to operationalize the WOSP criteria are given below.

6.3.1.1 Reliability of the Instrument. The Cronbach's Alpha values for the questionnaire statements on clarity for each of the eight factors are given in Table 6.1 below.

Table 6.1 Cronbach's Alpha for the Questionnaires on Clarity (N=60)

Factor	Mean	Variance	Std Dev	No of Variables	Alpha
Connectivity	56.62	55.87	7.47	13	.91
Extendibility	52.03	48.85	6.99	12	.90
Flexibility	61.22	62.41	7.90	14	.91
Functionality	64.78	82.14	9.06	15	.90
Privacy	49.53	32.93	5.74	11	.88
Reliability	50.78	67.16	8.19	12	.92
Security	58.85	45.01	6.71	13	.90
Usability	70.67	85.45	9.24	16	.93

Taking an alpha value of 0.7 to indicate that the instrument is reliable, the Cronbach's Alpha values for the questionnaires are higher than 0.8 for all eight factors. This indicates that the questionnaires are reliable for investigating the clarity of the performance factors.

6.3.1.2 Clarity of Statements. The results of the one-sample t-test on the rating of the clarity of the questionnaire statements are given in Table 6.2 below.

Table 6.2 One-Sample T-Test Results for Rating of Statements on Clarity (N-60)

Factor	Mean	Std. Deviation	t	df	Sig. (2-tailed)	Mean Difference
Connectivity	4.36	.58	18.26	59	.00	1.36
Functionality	4.32	.60	16.90	59	.00	1.32
Flexibility	4.37	.56	18.84	59	.00	1.37
Extendibility	4.34	.58	17.79	59	.00	1.34
Privacy	4.50	.52	22.27	59	.00	1.50
Reliability	4.23	.68	13.96	59	.00	1.23
Security	4.53	.52	22.90	59	.00	1.53
Usability	4.42	.58	19.01	59	.00	1.42

The above results show that for all factors, the t-values were significant at less than 0.01. This indicates that subjects found the statements used to operationalize the WOSP factors clear to understand.

6.3.2 Validity Aspect of the WOSP Criteria

The results of the instrument reliability and the one-sample t-test for responses regarding the validity of the statement used to operationalize the WOSP criteria are given below.

6.3.2.1 Reliability of the Instrument. The Cronbach's Alpha values for the questionnaire statements on validity for each of the eight factors are given in Table 6.3 below.

Table 6.3 Cronbach's Alpha for the Questionnaires on Validity (N=60)

Factor	Mean	Variance	Std Dev	No of Variables	Alpha
Connectivity	53.75	47.21	6.87	13	.84
Extendibility	50.27	47.35	6.88	12	.88
Flexibility	57.57	57.23	7.57	14	.85
Functionality	62.10	64.84	8.05	15	.84
Privacy	48.00	36.92	6.08	11	.86
Reliability	49.67	36.57	6.05	12	.80
Security	57.07	39.15	6.26	13	.83
Usability	67.35	75.52	8.69	16	.88

The Cronbach's alpha values for the questionnaire are higher than 0.8 for all eight factors. This indicates that the questionnaires were reliable for investigating the validity of the performance factors.

6.3.2.2 Validity of Statements. The results of the one-sample t-test on the rating of the validity of the questionnaire statements are given in Table 6.4 below.

Table 6.4 One-Sample T-Test Results for Rating of Statements on Validity (N-60)

Factor	Mean	Std. Deviation	t	df	Sig. (2-tailed)	Mean Difference
Connectivity	4.01	.47	16.63	59	.00	1.01
Functionality	4.12	.48	18.28	59	.00	1.12
Flexibility	3.93	.56	12.94	59	.00	0.93
Extendibility	3.91	.53	13.40	59	.00	0.91
Privacy	4.38	.46	23.01	59	.00	1.38
Reliability	4.21	.50	18.69	59	.00	1.21
Security	4.42	.43	25.66	59	.00	1.42
Usability	4.16	.53	16.84	59	.00	1.16

The above results show that the t-values are all significant at less than 0.01. This indicates that subjects found the statements used to operationalize the WOSP factors to be valid.

6.3.3 Importance Aspect of the WOSP Criteria

The results of the instrument reliability and the one-sample t-test for responses regarding the importance of the statement used to operationalize the WOSP criteria are given below.

6.3.3.1 Reliability of the Instrument. The Cronbach's Alpha values for the questionnaire statements on importance for each of the eight factors are given in Table 6.5 below.

Table 6.5 Cronbach's Alpha for the Questionnaires on Importance (N=60)

Factor	Mean	Variance	Std Dev	No of Variables	Alpha
Connectivity	52.18	37.68	6.14	13	.77
Extendibility	46.97	40.07	6.33	12	.80
Flexibility	55.02	60.83	7.80	14	.85
Functionality	61.82	50.73	7.12	15	.81
Privacy	48.18	26.12	5.11	11	.76
Reliability	50.48	35.95	5.80	12	.81
Security	57.43	30.96	5.56	13	.77
Usability	66.62	70.75	8.41	16	.86

The Cronbach's alpha values for the questionnaire are higher than 0.75 for all eight factors. This indicates that the questionnaires are reliable for investigating the importance of the performance factors.

6.3.3.2 Importance of Statements. The results of the one-sample t-test on the rating of the importance of the questionnaire statements are given in Table 6.6 below.

Table 6.6 One-Sample T-Test Results for Rating of Statements on Importance (N-60)

Factor	Mean	Std. Deviation	t	df	Sig. (2-tailed)	Mean Difference
Connectivity	4.14	.53	16.66	59	.000	1.14
Functionality	4.14	.54	16.44	59	.000	1.14
Flexibility	4.11	.54	15.94	59	.000	1.11
Extendibility	4.19	.57	16.05	59	.000	1.19
Privacy	4.36	.55	19.12	59	.000	1.36
Reliability	4.14	.50	17.51	59	.000	1.14
Security	4.39	.48	22.36	59	.000	1.39
Usability	4.21	.53	17.59	59	.000	1.21

The above results show that the t-values are all significant at less than 0.01. This indicates that subjects found the statements used to operationalize the WOSP factors to be important.

6.3.4 Implications for Hypotheses

The research question investigated in the statistical analysis carried out in this section is whether users understand each of the WOSP criterion goals as aspects of system performance. However, for the criteria to be helpful to users for the evaluation of given software, it must be possible to clearly operationalize the constructs. Furthermore, it must be possible to operationalize the construct in a manner that ensures content validity so that users can see the correspondence between the operationalization and the definition of the constructs. Finally, users must find that the operationalization is an important measure of the constructs. This gives rise to the following hypotheses. For each given criterion:

H1a: Users will in general understand the operationalization.

H1b: Users will in general consider the operationalization a valid measure.

H1c: Users will in general consider the operationalization to be important.

The results of the analysis show that the users agree to a statistically significant extent that the instruments used to operationalize the factor constructs are clear. H1a is therefore supported.

The results of the analysis also show that the users agree with a high degree of statistical significance that the instruments used to operationalize the factor constructs are valid. H1b is therefore supported.

Finally, the results of the analysis show that the users agree to a statistically significant extent that the factors as operationalized are important for the evaluation of software. H1c is therefore supported.

6.4 Preference for Using any of all the Eight WOSP Criteria

This section investigated whether users prefer to select from the set of all eight WOSP criteria when evaluating software in general. In other words the research issue is whether each criterion is considered important in general software selection. For this, preference for the use of WOSP factors in general software evaluation was compared to that for using just the duo of usability and functionality.

6.4.1 Preference for Choosing from all Eight WOSP Criteria

Table 6.7 below shows how many of the eight WOSP factors subjects preferred to use for general software evaluation. Of the 60 subjects, four did not fill in all the details required in the instrument. Their data was therefore excluded from analysis, leaving 56 subjects.

Table 6.7 Number of Factors Preferred for use in Software Evaluation

Number of factors preferred	Number of subjects	Percentage of subjects
8	38	67.9
7	4	7.1
6	1	1.8
5	4	7.1
4	5	8.9
3	4	7.1
2	0	0.0
Total	56	100.0

The results show that the overwhelming majority, almost 70%, of subjects prefer to use all eight WOSP criteria for evaluation purposes. All subjects prefer to use at least three of the factors. This means that subjects consider any 2-criteria method such as TAM to be inadequate for evaluation purposes, and that most of them do prefer the eight criteria available from WOSP.

6.4.2 WOSP Criteria vs. Functionality-Usability

Tables 6.8 and 6.9 below give various results for the comparison between the WOSP model and the common functionality-usability duo for general evaluation of software. Of the 60 subjects, 3 did not fill in all the details required in the instrument. Data from these 3 was therefore not used for analysis.

The statistics tabulated below in Table 6.8 show that the mean preference for the WOSP criteria by subjects was consistently higher than that for the functionality-usability (TAM) criteria for all three aspects of confidence, accuracy and completeness.

Table 6.8 Paired Samples Statistics (N= 57)

Aspect Compared		Mean	Std. Deviation	Std. Error Mean
Confidence	TAM	2.70	1.05	.14
	WOSP	4.63	.52	.07
Accuracy	TAM	2.72	1.01	.13
	WOSP	4.51	.57	.08
Completeness	TAM	2.40	.96	.13
	WOSP	4.56	.57	.08

The paired samples t-test results tabulated in Table 6.9 below show that, in addition to the WOSP criteria being consistently preferred to the TAM criteria for confidence, accuracy and completeness of evaluating software, the difference is statistically significant. The t value in all three cases is large, and is significant at less than 0.01.

Table 6.9 Paired Samples t-Test (N=57)

Comparison		Paired Differences		t	df	Sig. (2-tailed)
		Mean	Std. Deviation			
Confidence	TAM - WOSP	-1.93	1.25	-11.64	56	.00
Accuracy	TAM - WOSP	-1.79	1.25	-10.81	56	.00
Completeness	TAM - WOSP	-2.16	1.18	-13.84	56	.00

6.4.3 Frequency of the Criteria as a Critical Software Evaluation Factor

Tables 6.10 and 6.11 below give the results of statistical analysis of how frequently subjects would expect the various WOSP criteria to be a critical evaluation factor for many different types of software. Three of the subjects did not fill in all the details required in the questionnaire, and so their data was excluded from further analysis.

Table 6.10 Descriptive Statistics for Frequency of Criticality (N=57)

Factor	Mean	Std. Deviation
Usability	4.51	.66
Functionality	4.67	.51
Connectivity	3.91	1.04
Privacy	4.28	.98
Extendibility	3.65	.97
Security	4.81	.44
Reliability	4.79	.41
Flexibility	3.98	.86

Table 6.11 F-Values for Various Combinations of the WOSP Factors

No. of Factors	Factors	F-Value	Significance
All 8 WOSP Factors	Usability, Functionality, Security, Reliability, Privacy, Connectivity, Extendibility, Flexibility	12.73	0.00
5 most frequently critical	Usability, Functionality, Security, Reliability, Privacy	9.27	0.00
4 most frequently critical	Usability, Functionality, Security, Reliability	3.75	0.02
3 most frequently critical	Functionality, Security, Reliability	1.38	0.26

The descriptive statistics above show that the mean frequency of usage of the eight WOSP criteria is different for all eight performance factors. In addition, the F-test for the mean usage for all eight criteria shows that this difference is statistically significant. This indicates that the subjects recognize the factors as distinctly different. However, as the F-test is carried out successively for the five, four, and three factors that would be most frequently used the difference in frequency becomes correspondingly less significant as can be seen by the successively decreasing F-value. The difference in frequency of usage of the three most frequently used factors is not statistically significant. These results are noteworthy because they show that, while subjects recognize the factors as being different, they find the top three to be equivalent in terms of criticality for evaluating software. These are security, reliability, and functionality. It is also notable that reliability and security are rated higher in this regard than the more commonly encountered usability.

6.4.4 Implications for Hypotheses

The hypotheses deriving from the research question addressed in this section are the following:

H2: Users will prefer to use the WOSP criteria rather than just functionality and usability to evaluate software for an organization, and more specifically feel:

H2a: More confident in their choice using WOSP than just functionality and usability.

H2b: That the WOSP evaluation was more accurate than just functionality and usability.

H2c: That the WOSP evaluation was more complete than just functionality and usability.

H2d: Other WOSP evaluation criteria will be at least as frequently used as functionality and usability are on average.

The results show that the overwhelming majority of subjects prefer to use all eight WOSP criteria for software evaluation. Moreover, all subjects prefer to use at least 3 criteria for software evaluation rather than just functionality and usability. H2 was therefore supported by the results.

Moreover, the results show that subjects were, to a statistically significant extent, more confident in their choice using WOSP rather than just functionality and usability for performing software evaluation. This means that H2a was supported.

Similarly, subjects, to a statistically significant extent, felt that software evaluation was more accurate using WOSP criteria than when using just functionality and usability. Thus, H2b was supported by the experiment results.

In the case of completeness of evaluation, subjects, to a statistically significant extent, felt that software evaluation using WOSP criteria was more complete than when using only functionality and usability. H2c was therefore supported by the results.

Regarding frequency of usage of the WOSP criteria for software evaluation, statistically, subjects would use security and reliability as frequently as they would functionality. These three would be the most frequently used, and would be used more frequently than usability. This result supports H2d.

6.5 Users' Relative Rating of the WOSP Criteria for System Evaluation

This section discusses the results of the conjoint analysis which investigated how subjects rated the importance of WOSP factors relative to each other.

6.5.1 Analysis

Conjoint analysis calculates kendall's tau for holdouts as an indication of how consistent subjects are in ranking alternatives. For purposes of this research, a kendall's tau value of 0.4 and above was taken as an indication of subject consistency. Subject data that had a value lower than this was not used for aggregate analysis. This was the case for seven subjects. In addition, one subject had average importance values that were outliers, so the data was not used for aggregate analysis.

Only subject data that was statistically consistent as indicated by the kendall's tau value, and that did not have outliers was used. Results from 52 subjects were used for group aggregate analysis. In the aggregate analysis of conjoint analysis, one model is fitted to the aggregate results of all the responses. Table 6.12 gives the aggregate values of the individual subjects, as well as the subject group aggregate values. The two metrics by which the results of the experiment are given are adapted from (Bajaj 2000).

Ideally, if the performance factors are equally important to users, they should each have an average importance of 12.5%, since there are eight of them. In the first metric, the average importance values of the factors are presented. However, the mean is just a point estimate of the true value, whereas the confidence interval of the means is a better estimate of the true value of the parameter (Vining 1998).Therefore the 99% confidence interval is given for the average importance.

The second metric takes account of the possibility that the average importance could be biased by the extreme values of some subjects in the sample. Thus, the percentage of subjects in the sample that gave an average importance value equal to or greater than the expected 12.5% is given for each of the factors. The two metrics are given in Table 6.13. To investigate how the two metrics relate to each other, the correlation between the sets of data for the two is also given in the table. The 99% confidence interval of the average importance values is also shown graphically in Figure 6.1 below.

Table 6.12 Average Importance of WOSP Factors

	AVERAGE IMPORTANCE OF WOSP FACTORS PER SUBJECT								PEARSONS R		KENDALL'S TAU		HOLDOUTS KT	
S#	Extend.	Security	Connect.	Flex.	Reliab.	Usab.	Privacy	Funct.	Corr.	Sig	Corr.	Sig.	Corr.	Sig
1	6.67	34.20	6.96	3.77	5.22	8.12	26.09	8.99	0.94	0.00	0.83	0.00	0.87	0.01
2	3.49	38.41	5.40	2.54	5.71	2.54	34.92	6.98	0.98	0.00	0.94	0.00	0.60	0.05
3	5.16	8.02	8.88	29.80	10.60	13.75	18.34	5.44	0.89	0.00	0.79	0.00	0.73	0.02
4	7.40	6.11	3.86	6.43	2.25	27.01	13.18	33.76	0.92	0.00	0.78	0.00	0.69	0.03
5	2.64	30.50	7.92	7.33	6.74	8.21	23.17	13.49	0.90	0.00	0.77	0.00	0.60	0.05
7	11.20	32.77	1.68	1.68	18.77	12.89	5.60	15.41	0.96	0.00	0.88	0.00	0.60	0.05
9	1.29	11.61	0.97	3.87	10.32	30.65	17.74	23.55	0.92	0.00	0.81	0.00	1.00	0.00
10	10.08	20.65	2.77	2.02	14.86	17.13	20.40	12.09	0.94	0.00	0.81	0.00	0.87	0.01
12	11.29	8.24	14.59	15.53	13.18	12.94	14.12	10.12	0.93	0.00	0.87	0.00	0.87	0.01
13	12.30	22.40	3.55	7.38	19.13	7.92	21.31	6.01	0.94	0.00	0.85	0.00	0.87	0.01
14	6.14	55.29	6.83	5.80	6.14	5.80	7.85	6.14	0.99	0.00	0.95	0.00	1.00	0.00
16	19.09	2.42	5.76	7.27	36.97	6.97	5.45	16.06	0.95	0.00	0.83	0.00	0.47	0.09
17	5.29	13.53	3.53	17.35	8.53	13.82	6.47	31.47	0.93	0.00	0.85	0.00	0.47	0.09
18	2.29	31.81	7.45	2.87	10.89	22.06	18.62	4.01	0.94	0.00	0.88	0.00	0.87	0.01
19	14.36	19.14	12.59	1.76	10.83	17.38	21.41	2.52	0.95	0.00	0.88	0.00	1.00	0.00
20	6.97	9.41	18.12	6.62	2.79	40.77	8.36	6.97	0.85	0.00	0.66	0.00	0.60	0.05
21	3.57	26.53	5.36	3.32	16.07	10.97	21.43	12.76	0.96	0.00	0.70	0.00	1.00	0.00
22	13.90	21.34	17.62	12.66	9.68	2.23	4.22	18.36	0.96	0.00	0.90	0.00	0.73	0.02
23	6.65	14.10	12.77	4.52	10.37	29.79	11.17	10.64	0.92	0.00	0.80	0.00	0.73	0.02
24	5.22	34.62	21.70	6.87	2.20	11.26	12.09	6.04	0.99	0.00	0.94	0.00	0.73	0.02
26	13.89	27.78	5.83	3.06	26.39	13.33	3.06	6.67	0.95	0.00	0.83	0.00	0.60	0.05
28	3.24	35.10	24.78	4.72	6.19	1.47	11.50	12.98	0.94	0.00	0.83	0.00	0.60	0.05
30	4.29	35.66	12.87	6.97	9.65	9.38	10.46	10.72	0.96	0.00	0.88	0.00	0.87	0.01
31	0.00	69.23	0.00	0.00	7.69	0.00	23.08	0.00	1.00	0.00	1.00	0.00	1.00	0.00
32	13.85	10.80	17.14	11.27	13.15	8.69	12.44	12.68	0.92	0.00	0.85	0.00	0.87	0.01
33	11.05	34.28	9.35	4.82	4.82	10.76	18.41	6.52	0.92	0.00	0.76	0.00	0.47	0.09

Table 6.12 Average Importance of WOSP Factors (Continued)

	AVERAGE IMPORTANCE OF WOSP FACTORS PER SUBJECT								PEARSONS R		KENDALL'S TAU		HOLDOUTS KT	
S#	Extend.	Security	Connect.	Flex.	Reliab.	Usab.	Privacy	Funct.	Corr.	Sig	Corr.	Sig.	Corr.	Sig
34	9.05	8.07	13.20	7.33	21.76	12.71	10.51	17.36	0.93	0.00	0.84	0.00	0.87	0.01
35	12.21	11.50	14.55	10.09	14.79	10.80	15.26	10.80	0.92	0.00	0.85	0.00	0.87	0.01
38	5.96	42.01	18.81	5.96	8.15	12.23	5.02	1.88	0.93	0.00	0.80	0.00	0.73	0.02
39	6.08	14.55	9.26	9.52	6.88	25.66	5.82	22.22	0.93	0.00	0.81	0.00	0.47	0.091
40	6.65	13.57	27.70	3.32	20.22	22.16	3.32	3.05	0.96	0.00	0.89	0.00	0.60	0.05
41	7.86	12.47	5.69	12.74	19.24	17.07	15.99	8.94	0.86	0.00	0.81	0.00	0.60	0.05
43	6.93	20.27	20.27	6.40	16.00	6.40	9.07	14.67	0.91	0.00	0.79	0.00	0.87	0.01
44	5.16	17.12	3.53	7.88	17.66	23.37	13.32	11.96	0.89	0.00	0.75	0.00	0.47	0.09
45	3.37	23.31	4.60	10.74	36.20	4.60	15.34	1.84	0.99	0.00	0.85	0.00	0.60	0.05
46	6.83	27.02	1.86	13.98	17.08	6.21	19.88	7.14	0.86	0.00	0.71	0.00	0.73	0.02
47	18.11	8.16	26.28	11.22	6.12	5.87	13.27	10.97	0.93	0.00	0.86	0.00	0.73	0.02
48	10.39	13.63	15.70	12.70	12.93	11.09	14.09	9.47	0.97	0.00	0.90	0.00	0.87	0.01
50	7.67	4.11	8.77	6.03	20.27	13.15	19.73	20.27	0.89	0.00	0.81	0.00	1.00	0.00
51	4.00	18.86	18.29	13.43	11.71	20.86	9.14	3.71	0.86	0.00	0.71	0.00	0.60	0.05
54	3.72	2.58	23.21	9.46	31.52	16.05	8.02	5.44	0.95	0.00	0.84	0.00	0.60	0.05
56	11.26	21.72	16.62	10.72	6.43	12.60	10.99	9.65	0.85	0.00	0.83	0.00	0.87	0.01
57	5.78	22.17	8.43	9.40	17.83	9.88	10.12	16.39	0.95	0.00	0.85	0.00	0.60	0.05
59	13.33	5.64	14.36	13.08	16.92	14.62	7.95	14.10	0.88	0.00	0.72	0.00	0.47	0.09
60	4.09	21.38	4.40	6.29	42.77	5.03	5.66	10.38	0.99	0.00	0.92	0.00	0.87	0.01
62	10.74	14.05	26.72	4.68	3.31	5.51	11.85	23.14	0.93	0.00	0.86	0.00	0.87	0.01
63	9.60	13.35	15.22	11.01	15.93	8.43	12.65	13.82	0.93	0.00	0.83	0.00	0.87	0.01
65	13.30	16.37	11.51	17.39	12.28	12.02	10.74	6.39	0.86	0.00	0.78	0.00	0.73	0.02
66	6.32	28.42	9.21	6.05	21.32	12.37	0.53	15.79	0.97	0.00	0.88	0.00	0.73	0.02
69	1.94	11.67	8.06	10.56	16.39	19.72	11.11	20.56	0.87	0.00	0.88	0.00	1.00	0.00
70	9.17	21.89	7.10	3.25	22.49	7.10	26.33	2.66	0.89	0.00	0.77	0.00	0.97	0.00
71	14.78	22.69	17.94	1.58	5.80	10.55	6.86	19.79	0.94	0.00	0.85	0.00	0.47	0.09
Av.	7.99	20.97	11.34	7.98	14.06	12.92	13.14	11.59	0.95	0.00	0.87	0.00	0.60	0.05

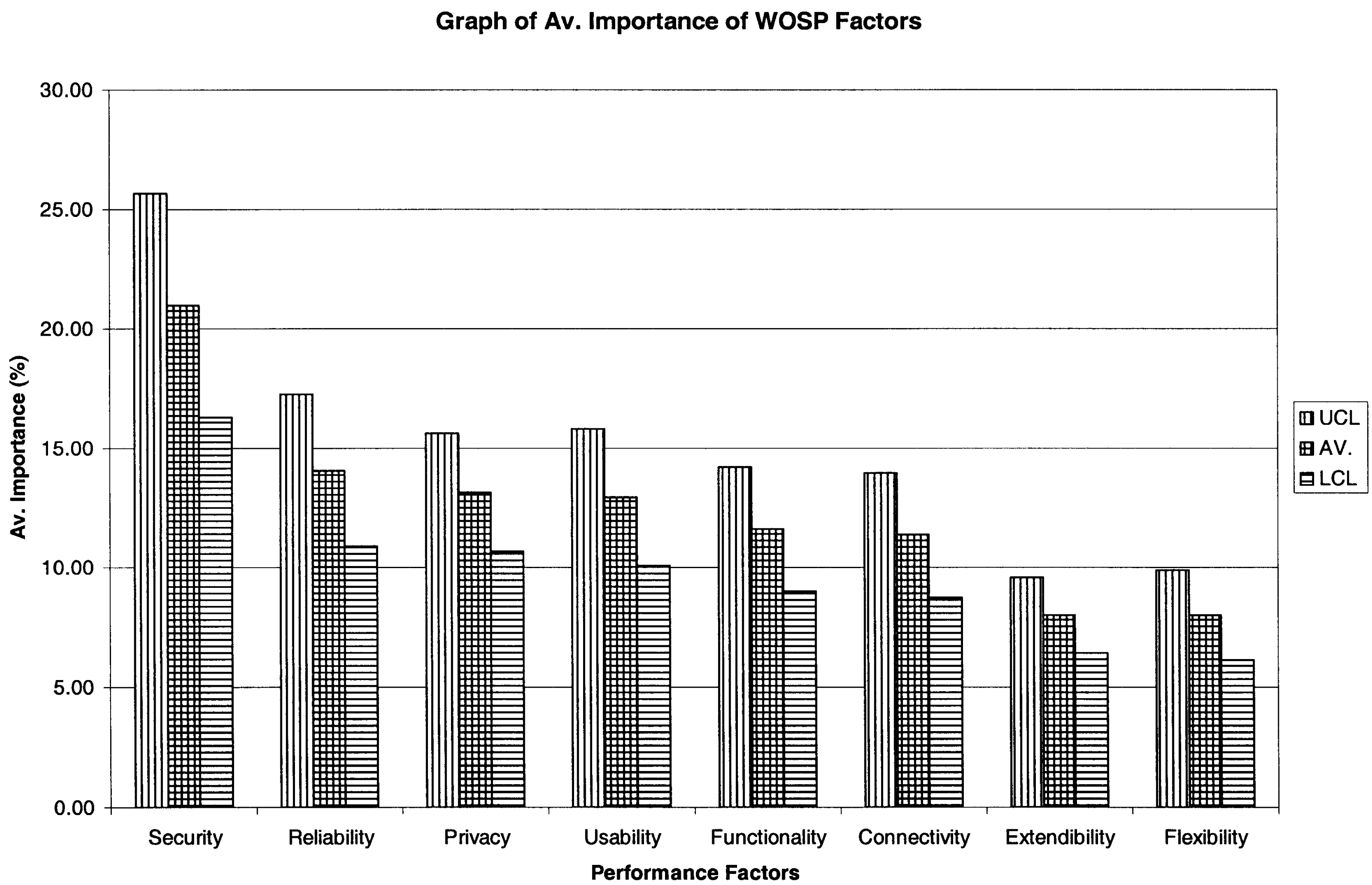

Figure 6.1 Graph of average importance of WOSP factors.

Table 6.13 Correlation between Average Importance of WOSP Criteria and Percentage of Subjects giving Criteria an Average Importance Higher than 12.5%

Factor	Std. Dev.	Av. Importance	99% Confidence	%age above 12.5%
Security	4.40	20.97	16.28 - 25.66	73.08
Reliability	13.13	14.06	10.86 - 17.26	51.92
Privacy	7.31	13.14	10.65 - 15.63	48.08
Usability	5.25	12.92	10.04 -15.80	46.15
Functionality	8.95	11.59	8.99 - 14.19	40.38
Connectivity	8.06	11.34	8.73 - 13.95	44.23
Extendibility	6.98	7.99	6.42 - 9.56	19.23
Flexibility	7.27	7.98	6.11 - 9.85	19.23
Correlation between av. Importance and percentage of subjects giving factor an importance of at least 12.5%				0.97

6.5.2 Implications for Hypotheses

If WOSP criteria in fact improve on functionality and usability for software evaluation, some of the other factors should play a significant role in technology acceptance. If they rate the same as, or higher than, the WOSP functionality and usability which correspond to the very commonly encountered and applied constructs of usefulness and ease of use from TAM, then the WOSP additional criteria are useful. However, if the top evaluation criteria are functionality and usability, and other factors are much lower, then the WOSP contribution is minimal. This leads to the following hypothesis:

H3a: in the WOSP evaluation, performance criteria other than functionality and usability will contribute significantly, and so be seen as important, to technology acceptance.

Ideally, if the performance factors are equally important to users, all the eight performance factors should be perceived to have an equivalent effect on the performance of software. Thus, if the total performance of software is taken as 100%, they should each account for 12.5% of the evaluation decision, since there are eight of them. However, since the WOSP model predicts that the extent to which the individual performance factors are important depends on both the system and the environment in which it is operating, the factors are all not necessarily of equal importance. This leads to the following hypothesis:

H3b: all eight factors are not considered to be equally important by users.

For the purposes of evaluating the results, it is will be assumed that if the expected value of 12.5% is either contained within or lower thanthe confidence interval of a factor, that factor is significant in the subject's performance evaluation model.

Security: The average importance for security was the highest among the eight factors. The expected average importance was lower than the confidence interval. This factor also had the highest percentage of subjects with an average importance greater than 12.5%.

Reliability: The average importance for reliability was the second highest of the eight factors. The expected average importance was within the confidence interval. This factor also had the second highest percentage of subjects giving it an average importance greater than 12.5%.

Privacy: This factor had the third highest average importance for the subject group. The expected average importance was within the confidence interval. Privacy had the third highest percentage of subjects giving it an average importance greater than 12.5%.

Usability: Usability had the fourth highest subject group average importance. Also, the expected average importance was within the confidence interval. Usability had the fourth highest percentage of subjects giving it an average importance greater than 12.5%.

Functionality: This performance factor had the fifth highest subject group average importance. The expected average importance fell within the confidence interval. The factor had the sixth highest percentage of subjects giving it an average importance greater than 12.5%.

Connectivity: This factor had the sixth highest subject group average importance. The expected average importance was within the confidence interval. Connectivity ranked fifth in percentage of subjects giving it an average importance greater than 12.5%.

Extendibility: Extendibility had the seventh highest and second lowest subject group average importance. The expected average importance falls beyond the confidence interval. It also had the seventh highest, or second lowest percentage of subjects giving it an average importance of 12.5% or higher.

Flexibility: This factor had the eighth and lowest average importance from the subject group. The expected average importance is beyond the confidence interval. This factor also had the lowest percentage of subjects giving it an average importance of 12.5% or greater.

Furthermore, the correlation between the two metrics of average importance and percentage of subjects giving a rating of at least 12.5% is 0.97, which is very high, and is an indication of the consistency of the results.

The results therefore support H3a since security, reliability and privacy are all ranked as more important than both usability and functionality.

The results support H3b, since security is rated as more important than the other factors, while extendibility and flexibility are rated as less important than the other factors.

6.6 Checking the Effect of Operationalization on Relative Importance of Criteria

This section deals with the issue of whether the manner in which the performance criteria were operationalized may have had any effect on the findings of the relative importance of these criteria.

6.6.1 Analysis

To investigate the possible effect of the operationalization statements in research question 1 on the relative importance values of the criteria in research question 3, correlations between the average importance of the factors as obtained from conjoint analysis in research question 3 and the rating of clarity, validity and importance by the subjects obtained from the questionnaires used for research question 1 was done. The results are given in Table 6.14 below.

Table 6.14 Correlation between Average Importance and Clarity, Validity, and Importance Ratings

Factor	Av. Importance	Std. Dev.	99% Confidence	Clarity Ratings	Validity Ratings	Importance Ratings
Security	20.97	4.40	16.28 - 25.66	58.85	57.07	57.43
Reliability	14.06	13.13	10.86 - 17.26	50.78	49.67	50.48
Privacy	13.14	7.31	10.65 - 15.63	49.53	48.00	48.18
Usability	12.92	5.25	10.04 -15.80	70.67	67.35	66.62
Functionality	11.59	8.95	8.99 - 14.19	64.78	62.10	61.82
Connectivity	11.34	8.06	8.73 - 13.95	56.62	53.75	52.18
Extendibility	7.99	6.98	6.42 - 9.56	52.03	50.27	46.97
Flexibility	7.98	7.27	6.11 - 9.85	61.22	57.57	55.02
Correlation between av. Importance and the clarity, validity and importance ratings. The p-values are in brackets				0.02 (0.97)	0.08 (0.86)	0.24 (0.57)

6.6.2 Implications for Hypotheses

It is logical to suppose that in the event that the clarity, validity, and importance of the operationalizations of the eight criteria are significantly different, then the clarity, validity, and importance of the operationalizations of the eight criteria would have an effect on the importance that users attach to each of the criteria. This leads to the following set of hypotheses:

H4a: Criteria operationalizations that are clearer will lead to higher importance being attached to the respective criteria.

H4b: Criteria operationalizations that are more valid will lead to higher importance being attached to the respective criteria.

H4c: Criteria operationalizations that are more important will lead to higher importance being attached to the respective criteria.

The results indicate that there is little correlation between the perception of both clarity and validity and the average importance of the factors. In other words the statements might be perfectly clear, and may be deemed to be a perfectly valid operationalization of the factor. The statements may also be an important consideration of the criterion. This does not however mean that the factor will be rated correspondingly highly in importance.

It should be noted that the rating was more a comparison of the statements of a given performance factor with each other, rather than a comparison of the statements for the various factors with each other. The correlation with average importance is therefore more of an indication of what, if any, the relative relationship of clarity, validity, and importance of the operationalization statements – when used for priming for example -

might have with the actual importance assigned to the factors, or whether these statements might have an effect on the average importance of the criteria.

6.7 Summary of Hypotheses

The outcome of the hypotheses proposed in the research is summarized below:

H1a Users will in general understand the operationalizations

SUPPORTED

H1b Users will in general consider the operationalizations a valid measure of the constructs

SUPPORTED

H1c Users will in general consider the operationalizations important for software selection

SUPPORTED

H2 Users will prefer to use the WOSP criteria rather than just functionality and usability

SUPPORTED

H2a In comparison to functionality-usability, users will feel more confident in their choice In comparison to functionality-usability, users after using the WOSP criteria

SUPPORTED

H2b In comparison to functionality-usability, users feel the WOSP evaluation is more accurate

SUPPORTED

H2c In comparison to functionality-usability, users feel that the WOSP evaluation is more complete

SUPPORTED

H2d In comparison to functionality-usability, users use other WOSP evaluation criteria at least as frequently

SUPPORTED

H3a Performance criteria other than functionality and usability will be significant

SUPPORTED

H3b All 8 performance criteria will be considered equally important by users

SUPPORTED

H4a Criteria operationalizations that are clearer will lead to higher importance being attached to the respective criteria

NOT SUPPORTED

H4b Criteria operationalizations that are more valid will lead to higher importance being attached to the respective criteria

NOT SUPPORTED

H4c Criteria operationalizations that are more important will lead to higher importance being attached to the respective criteria

NOT SUPPORTED

6.8 Discussion

It is surprising that functionality and usability, which TAM predicts as primary user selection criteria, rank fourth and fifth, respectively. While TAM has been proven to be valid, and may be the primary user acceptance theory, it seems not to represent the primary user acceptance factors for social-technical systems.

It is also notable that reliability has a distinctly different weight from security, suggesting that software users distinguish these concepts from each other. While some technology acceptance models present reliability as an aspect of security (Laprie 1992), others see it as "the ability of a system to resist attack" (Littlewood et al. 1993). This approach explains why mechanisms that increase fault-tolerance (reliability) can reduce system security (which is illogical if reliability is an aspect of security). Recent models reclaim the reliability/security distinction, as the first is based on provision of service, while the second is based on denial of service (Jonsson 1998).

Similarly, while security has been described in various IS literature as including confidentiality (privacy) as an attribute (Defence 1985; European Commission 1993), the results suggest that security and privacy are distinct in the minds of users. Likewise, while flexibility has been suggested to include scalability (an aspect of extendibility) and connectivity (Knoll and Jarvenpaa 1994), the results imply that flexibility, extendibility and connectivity are recognized as distinct concepts by users.

Despite the recent IS interest in flexibility under terms like mobility and agility, this study found flexibility had the lowest relative importance. This may have been because flexibility is not very relevant to browser software, or because many users do not fully understand or appreciate its value. Or it may be because although flexibility is very important to system designers and programmers, it is not so important to users.

Generally, of the eight factors the WOSP model presents, extendibility, flexibility and connectivity are the least known. More familiar factors like security, usability and reliability may be rated as more important, following the availability heuristic (Kahneman, Slovic and Tversky 1982). However, as users become more knowledgeable, and as use of social-technical systems becomes more established, the appreciation of the various performance factors may change, with an accompanying change in the rating of software acceptance factors.

This research found that the WOSP model provides a more comprehensive set of criteria than any other current model does for the user evaluation of the intrinsic technical performance factors of complex socio-technical information systems such as browsers.

6.9 Research Limitations

The research focused on the applicability of the WOSP criteria by users in assessing how information systems perform. It did not investigate the design aspect of the WOSP model, for example resolving the tensions that arise in fulfilling the various design requirements, or even how to go about fulfilling the requirements. While this is not a limitation of this research, it is useful to bear in mind such other related aspects of the model. Several limitations of this research were identified. The following is a discussion of some of the main ones associated with the study.

The model does not have a time dimension as one of the parameters of interest to users when evaluating the performance of a system. Given the intensely competitive nature of business in today's commercial activities for example, time is of major consequence in most transactions of the business world. Similarly, the cost of procuring the information systems was not considered. Thus the WOSP criteria were investigated without taking into consideration their impact on the cost of an information system. In reality however, cost is a very important consideration for decision makers in organizations.

In theory, the WOSP criteria are applicable to all applications. However, the prominence of the various criteria varies from system to system and also depends on the context in which the system is used. The results of this research validated the WOSP criteria with one specific application – the browser. It would be very useful for similar validation with other applications to verify the generalizability of the criteria.

While the WOSP model criteria for intrinsic system qualities are more comprehensive than any other model so far, they are not necessarily complete. For example, privacy and connectivity bring into consideration social processes. Consideration of other social contexts in which socio-technical systems operate, such as trust, may lead to other constructs and criteria that have not been included in the current WOSP model.

In this research, subjects were asked to take the role of managers in performing certain tasks. Most of the students who participated in the research are full time employees who pursue their graduate studies part time. While this is not considered a major limitation, especially given that all had extensive experience in the experimental software, there is a possibility of a difference in the decision making process with fulltime managers that should be borne in mind. For purposes of validating the results, it would be useful to collect data from subjects who are actually fulltime managers.

CHAPTER 7

FUTURE RESEARCH AND CONTRIBUTION OF WORK

As was pointed out in the introduction chapter, information technology has become a primary survival factor in many business organizations. Its being considered a critical success factor has led to these organizations investing heavily in technology. In addition, information technology has come to play a critical role in other areas such as health care, education, mass transportation, and security, for example.

Despite this, however, many decision-makers have a difficult time justifying investment in IT because they are not sure that the systems procured meet user expectations for performance in their organizations. A systematic way for users to evaluate performance of information systems of interest would go a very long way in helping alleviate these investment concerns. Indeed, such a means would be of great benefit to society at large, as it would be most helpful for any prospective users of information systems, no matter the context, to be able to better assess the performance of an information system he or she might be interested in. This chapter summarizes the contributions of this research, and points out possible areas for future research. It is broadly categorized into theoretical contributions, application in industry and organizational context.

7.1 Theoretical Aspect

While many models and theories have been put forward to explain the usage of information systems by individuals, hitherto none of the established ones, with the exception of the Technology Acceptance Model, has sought to explain this user behavior purely in terms of the intrinsic characteristics of the information system. While providing very useful insight into users' assessment of information systems, they leave unclear the specific role and extent of influence of the factors in the three broad categories of individual characteristics, technological characteristics, and organizational characteristics that have a hand in the evaluation of information systems. In this regard, the WOSP model is a welcome addition to the TAM in helping clarify the role of the intrinsic qualities of information systems on user evaluation.

In this research, the WOSP model was found to be the most comprehensive model in terms of inclusion of software quality factors, as it encompasses other existing models. The eight factors for system performance proposed in the WOSP model were confirmed to be valid intrinsic software characteristics, well-documented in information systems literature. The results of this study suggest that the WOSP model adds to TAM factors that are well recognized in systems requirements literature.

The experimental software for this research was a browser. Although browsers are quite complex organizational level applications, they are not the only example, and not necessarily the most complex applications that are also organizational in scope. It would be useful to repeat this experiment with other software of comparable complexity (Venkatesh et al. 2003). This might give a fuller picture of how the factors affect software selection. Also, such further studies may define different factor configurations for other types of software.

Moreover, it would be useful to generalize the statements used to anchor this browser evaluation to create an eight factor equivalent of the TAM questionnaire. The statements for each performance criterion would then constitute a scale. These scales would then be rated against each other, and the rating compared to the average importance from a conjoint analysis of various information systems. If there is a consistently close correlation for different software, these scales would be a simple and very useful instrument for users to evaluate the performance of information systems.

7.2 Industry Application

The WOSP model gives a straightforward way for users to indicate their software preferences to designers of information systems. In conjunction with an analysis method such as Conjoint Analysis, which analyses at the individual and aggregate level, the WOSP model can facilitate the following product development functions (Hair et al. 1995) briefly discussed below.

7.2.1 Segmentation

The analysis could be used to segment the users of a given application of interest according to the importance they attach to each of the eight performance factors. This could go a long way in helping alleviate some of the major managerial problems like resistance to the adoption of certain applications by certain individuals who perceive the application as lacking some of the qualities that they consider essential for their acceptance. For instance, privacy might not be as much of an issue to most high school and undergraduate students as connectivity, as they may be relatively unaware of the consequences of lack of privacy which is to a large extent shaped by the social environment, while it might be critical in some work environments.

7.2.2 Marketing Information

For developers and vendors of applications, information obtained on the relative importance of the performance factors by various target user groups could be combined with information on the cost of providing the various factors, and even the cost of providing various levels of each factor, at the development stage. This could give useful insight on the profitability margin of providing various groups with their ideal applications, and how to optimize the development of applications.

7.2.3 Simulation

A related use of conjoint models is its application in simulation. This would involve the three steps below.

- Estimation and validation of conjoint models for subjects drawn from a population of interest.
- Selection of stimuli for testing that would be based on an issue for research in the software of interest.
- Simulation of the choices by the subjects for the selected stimuli to predict what the evaluation of the application by the subjects would be.

This simulation procedure would be particularly useful in researching new information system products, as it would be possible to have a good idea of what a particular target population might prefer in a system with relatively very little expenditure of resources. And this would be a helpful step to perform in conjunction with, or as a prelude to, the design of a prototype.

7.3 Organizational Context

The context within which this research was done is organizational. In many organizations today, considerations of gender, age and experience are becoming increasingly important in a bid to provide equitable opportunities to, and maximize the productivity of, all constituent workplace populations. With the proliferation of internet and communication technologies in the workplace, it is pertinent to investigate the consequences of these human factors on the acceptance and use of the information systems in organizations. These factors are now discussed below within the framework of the system performance factors of the WOSP model.

7.3.1 Functionality

Available research indicates that men in general are highly task-oriented (Minton and Schneider 1980). This suggests that system qualities that facilitate enhanced functionality would be of particular interest to men, or that men would tend to rate functionality as being of greater relative importance to men than women would. A further development of this is that gender roles and the associated socialization processes that are reinforced from birth, rather than the actual biological gender is what is of consequence (Lubinksi et al. 1983). UTAUT's performance expectancy is defined as the degree to which an individual believes that using an information system will help him or her attain gains in job performance (Venkatesh et al 2003). It includes perceived usefulness and is closely related to WOSP's functionality factor. Indeed gender was found to be a factor moderating performance expectancy (Venkatesh et al 2003). This justifies further research with regards to WOSP's functionality factor.

Research has also shown that younger workers may be more concerned about 'extrinsic' rewards than their older counterparts (Hall and Mansfield 1975, Porter 1963). Similarly, age was found to moderate the extent to which performance expectancy affected behavioral intention to use an information system, with this being more salient to younger workers (Venkatesh et al 2003). However, other studies suggest that studying the effect of gender without taking age into account may be misleading (Levy 1988). One reason for this cited in the literature is that the importance of job related factors to users may vary significantly with corresponding changes in family-oriented and other values over one's adult life (Barnett and Marshall 1991). Again, this warrants further research in the context of WOSP

7.3.2 Usability

WOSP's usability factor is closely related to UTAUT's effort expectancy which is defined as the degree of ease associated with the use of the system (Venkatesh et al 2003). Previous research has suggested that this factor is of more significance to women than to men (Bern and Allen 1974, Bozionelos 1996). As with performance expectancy, however, this may have more to do with gender roles, rather than the actual biological gender.

Moreover, prior research suggests that age is associated with increased difficulty in processing complex stimuli as well as a decreased ability to give tasks appropriate attention (Plude and Hoyer 1985). Furthermore, past research has given support to the view that constructs with a bearing on effort expectancy will exert a stronger influence on intention for women and for older workers. And actually, this notion on effort expectancy was supported by (Venkatesh et al 2003), with the construct being more important to women, and particularly older women. This gives grounds for similar research on WOSP's usability factor.

7.3.3 Privacy and Connectivity

Privacy is defined in the WOSP model as the ability to limit unwanted information disclosure. It is indicative of how much an individual cares about how others in society perceive them. Similarly, connectivity as defined in the WOSP model is the ability to exchange information with other systems, and can also be described using such words as sociability, connectedness, networking, and communication. In this regard, both WOSP constructs can be seen as being related to the social influence construct under UTAUT, which is defined as the degree to which an individual perceives that important others believe he or she should use the new system.

Research theory suggests that women tend to be more sensitive to the opinions others hold about them (Miller 1976, Venkatesh et al. 2000). Such being the case, factors with strong connotations of social influence, such as privacy and connectivity, may be of more importance to women in their evaluation of information systems. Moreover, with time, people tend to accumulate more and more information about themselves, the exposure of which that they would want to have more control over. Age and gender were found to moderate the effect of UTAUT's social influence, with the construct being more significant for women, and more so older women. The effect of gender, age, and experience on the perceived importance of privacy and connectivity is therefore a possible area of further research.

7.3.4 Extendibility

In the WOSP model, extendibility is defined as the ability to make use of third party programs and data. This appears related to UTAUT's construct of facilitating conditions, which is defined as the degree to which an individual believes that an organizational and technical infrastructure exists to support use of the system, and which encompasses notions such as a system's compatibility with other systems used.

It has been noted in the literature that the consequence of compatibility is expected to increase with experience, as users find ways of help and support, perhaps even other software that it would be useful to have (Bergeron et al. 1990). Furthermore, organizational psychologists have made the observation that older workers attach more importance to receiving help and assistance on a task (Hall and Mansfield 1975). This is more so given the cognitive and physical limitations that come with age. In the context of information systems, this help might be in the form of third party programs, for example.

In the case of the facilitating conditions construct of UTAUT, age and experience were found to exert a moderating influence, with the construct being more important to older workers, and with increasing experience. It would therefore be worthwhile to investigate the effect of age and experience on the relative importance users attach to extendibility in the WOSP model.

7.3.5 Security, Flexibility, Reliability

No literature was reviewed that explicitly or implicitly discussed the effect of age, gender or experience on the above three WOSP constructs or any related concepts. However, a suggestion common from the definitions of these constructs is that they would likely become more valued with experience and use of information system. One typically becomes more appreciative of security after several mishaps due to a lack of it. Similarly, one comes to value reliability with time, as it connotes predictability which is essential for planning and for the organized pursuit of objectives. Flexibility, which has to do with the ability to change to suit different situations, proves valuable over time, when one has to deal with unexpected circumstances, and realizes that these are to be expected. It is reasonable to expect that age and gender would be of some consequence to these three as well, given the suggestions to this effect in the literature on the other five factors. This is therefore a possibility for future research as well.

7.4 Conclusion

Information and communication technology has had a profound and irreversible impact on society in myriad aspects, from personal to public, and from household to global. Information systems have played a most significant role in this revolution. As society becomes ever more reliant on this technology in all facets of life as described in the introductory chapter, information systems will likely evolve into correspondingly more complex socio-technical systems.

In order to optimize the usage of these systems, it will be necessary to understand how their intrinsic qualities impact all potential users from diverse cultures and other constituencies. Examples of such systems currently existing, whose full potential would be greatly enhanced by such an understanding, include the following:

- Voter systems.
- E-commerce and banking systems.
- Health delivery systems such as telemedicine.
- Emergency response systems.
- Online educational systems.
- Security systems.

The research presented here suggests that the WOSP model would be most useful indeed for carrying out the necessary comprehensive evaluation on these and other similar systems.

APPENDIX A

STATEMENTS FOR OPERATIONALIZING CRITERIA FOR THE ASSESSMENT OF THE BROWSERS

This appendix contains the statements that were used to operationalize the eight WOSP criteria for the assessment of browsers in Research Question 1.

<u>WEB BROWSER PERFORMANCE STATEMENTS TASK</u>

Subject Details

Please complete the following:

1. Subject ID: _____
2. Gender: M / F
3. Approximately how many years/months have you used a browser? ____ years ___ months
4. What are the three main reasons you use a browser?

.....

5. <u>In the last 6 months</u>, how many hours per week, on average, have you used a browser to access and search the Internet? ___________ hours per week, on average

Performance Statement Ratings

The performance statements are organized into eight performance dimensions, each on a new page. For each of the eight browser performance dimensions:

1. *Read the definition and similar terms that are given on the top left of every page.*
2. *Read and rate the browser statements.* For each say if you feel it is:
   a. **Clear?** Do you understand what it means?
   b. **Valid**? Is it a valid statement of the performance dimension definition?
   c. **Important?** Is it important when selecting a browser?

   Answer 1 – 5 as follows:

   1= Strongly Disagree, 2= Disagree, 3= In the middle, 4=Agree, 5= Strongly Agree
3. *Make comments.* Suggest things left out, give reasons for grading statements low.

Flexibility: The ability to change itself to fit different situations. *Similar terms*: adaptability, portability, customizability, plasticity, agility	1= Strongly Disagree, 2= Disagree, 3= In the middle, 4=Agree, 5= Strongly Agree		
	Is the statement		
Browser Statements	**Clear?**	**Valid?** by the definition	**Important?** in browser selection
It runs on all our computers and operating systems.			
It has a preferences "control panel" to change browser settings.			
Its settings can be changed easily and quickly.			
Graphics, sound and video settings can be adjusted so it works OK on low-end machines.			
It can work in different languages, by clicking buttons with flags of different countries.			
I use it for many types of tasks, and can usually adapt it for new and unexpected tasks.			
It can be used in any organizational environment without much modification.			
It analyses download rates by site, so I can select responsive sites.			
I can change its "skin", or colors and appearance, easily.			
It is easily changed to fit disability needs e.g. larger text or graphics for those with poor vision.			
It can save preference settings under a name, so I can reload my preferences easily.			
It adapts to user response patterns, so frequent user menu choices go to the top of the menu.			
The user can choose “Novice” (simple) or “Advanced” menus and dialogues			
I can change the toolbar to include any buttons in any order.			

Please comment on your ratings:

Reliability: The ability to continue working despite errors or problems, or to quickly recover if failure occurs. *Similar terms*: stability, dependability, recovery, durability, ruggedness	1= Strongly Disagree, 2= Disagree, 3= In the middle, 4=Agree, 5= Strongly Agree		
	Is the statement		
Browser Statements	**Clear?**	**Valid?** by the definition	**Important?** in browser selection
The number of errors while using it is low, or lower than average.			
I am never stuck, because there is always another way to do the same thing, e.g. mouse click or keyboard.			
Even if I give it many commands quickly one after the other, it still works well.			
Even if I multi-task, and do many things at once, it still works well.			
It never breaks down or "hangs" (fails to respond), even if I use it for a long time.			
If one part of the browser fails, like a plug-in, the entire browser does not crash.			
If a web page part cannot be displayed, it offers other ways to display it successfully that usually work.			
The software gives useful error messages that help users do the right thing the next time.			
It can check to see if my "Favorites" links still work, and lets me delete those that do not			
If the browser fails, the problem is usually easy to repair and of short duration.			
It can suppress web site errors that repeat, so they do not interfere with browsing.			
It can backup its data, like my favorites list, to a single file that can be reloaded if there is a problem			

Please comment on your ratings:

Security: The ability to defend against hostile attack, including unauthorized entry, change, damage, hurt or takeover. *Similar terms*: defendability, protectiveness, resistance to attack.	1= Strongly Disagree, 2= Disagree, 3= In the middle, 4=Agree, 5= Strongly Agree		
	Is the statement		
Browser Statements	**Clear?**	**Valid?** by the definition	**Important?** in browser selection
When a file is downloaded to the hard drive, it is checked for viruses before use.			
It can be set to avoid pornography sites or other unwanted content.			
Its security features are easy to use and are effective.			
I can select different degrees of security, to trade-off security and other features.			
It tells me what the current security threat level is.			
It gives me a choice to allow java applets and ActiveX programs to run or not.			
It advises if a web site tries to use my connection to make a phone call, and can prevent it.			
It tells me if a web site attempts to change my browser settings, like my home page, and can prevent it.			
It can detect and prevent popup ads.			
It can detect and prevent spyware from installing.			
It keeps usage logs, so I know if someone other than me used the system.			
It can stop new browser windows opening without my permission.			
It has a log-on/password feature so I can prevent others from using my browser.			

Please comment on your ratings:

Extendibility: The ability to make use of third party programs and data. *Similar terms*: openness, compatibility, scalability, open standards, tool use.	1= Strongly Disagree, 2= Disagree, 3= In the middle, 4=Agree, 5= Strongly Agree		
	Is the statement		
Browser Statements	**Clear?**	**Valid?** by the definition	**Important?** in browser selection
It can save web site sounds to my computer in wave, midi and most other formats.			
It lets me view a web site's source code with my preferred editor/viewer.			
The email button starts the email system of my choice, not just the given browser email.			
It works with all third party multimedia tools, like real-media player and flash.			
I can copy graphics from a web site to my computer.			
It can install plug-ins for the latest Internet music, radio, TV and video.			
The browser is open-source, so trusted third parties can develop useful add-in modules.			
It follows all World Wide Web source code and data standards, e.g. unicode.			
It can handle graphics, sound and video in a wide variety of different formats.			
Its software architecture conforms to accepted programming standards.			
I can clip and paste from the browser into any company software application.			
It can import and export favorites lists from and to other browsers.			

Please comment on your ratings:

Privacy: The ability to limit unwanted information disclosure. *Similar terms*: confidentiality, secrecy, camouflage, stealth, opaqueness.	1= Strongly Disagree, 2= Disagree, 3= In the middle, 4=Agree, 5= Strongly Agree		
	Is the statement		
Browser Statements	**Clear?**	**Valid?** by the definition	**Important?** in browser selection
It stops web sites from reading my browsing history.			
It stops web sites from getting my name or email from my computer's data.			
It lets me decide if a web site stores a cookie or not.			
It lets me review web site cookies, shows who stored it and when, and lets me delete those I don't want.			
It lets me browse the Internet anonymously if I want to.			
When browsing, I can specify what information to release, and what to keep private from a website.			
Any sensitive information I give the browser, like logon passwords, is encrypted, so others can't see it.			
Password information always shows as asterisks, so others cannot look over my shoulder to see them.			
I can click a button to hide the browser page, making it disappear, and then click again to bring it back.			
It can clear all my browsing history at the click of a button.			
It prevents online tracking, spam harvesting, and snoops while I am browsing.			

Please comment on your ratings:

Connectivity: The ability to exchange information with other systems. *Similar terms*: interactivity, sociability, connectedness, networking, communication.	1= Strongly Disagree, 2= Disagree, 3= In the middle, 4=Agree, 5= Strongly Agree		
	Is the statement		
Browser Statements	**Clear?**	**Valid?** by the definition	**Important?** in browser selection
It can monitor a web site in the background, and advise me if anything on it changes or is added.			
It can download many files at once, without conflicts.			
When downloading it gives useful information, like the estimated time to complete the download.			
If a download is stopped, or fails for any reason, it can be restarted again later from where it left off, saving a lot of time.			
It has a utility to "tweak" Internet connection settings for faster access			
It tells me if new browser software version upgrades are available to download.			
It lets me browse the Internet with others, as a group, so we can all look at the same sites.			
Previously downloaded graphics and web files are stored and re-used, so familiar sites load very quickly			
When loading a very "rich" web site, it tells you what percentage has loaded, and how long to go			
It can prevent my connection from timing out.			
It can download data sites in the "background", without affecting my current browsing at all.			
It can generate an email message to anyone in my address book with a link to a web site that I like			
It gives access to other ways of communicating, like Telnet, Ftp, email and chat.			

Please comment on your ratings:

Functionality: The ability to move to web sites and display their information. *Similar terms*: effectiveness, capability, usefulness, effectiveness, power, strength.	1= Strongly Disagree, 2= Disagree, 3= In the middle, 4=Agree, 5= Strongly Agree		
	Is the statement		
Browser Statements	**Clear?**	**Valid?** by the definition	**Important?** in browser selection
This browser gets me where I want to go quickly			
Using it improves my work performance.			
It remembers where I was last, and returns there when I reopen the browser later.			
It can open many websites at once, and I can jump between them by clicking on a tab or button.			
Using it enhances my effectiveness on the job.			
The back and forward buttons are drop-down lists, so I can navigate my browser history			
Displayed web pages look very good on the screen with this browser.			
It can extract all the links, emails addresses, graphics or sounds in a web site, giving an alphabetic list			
Using it in my work increases my productivity.			
It can open and display web files from my hard drive			
It can download web sites to the hard drive, for browsing later while offline			
The Favorites list lets me jump directly to my favorite sites			
I can organize the Favorites list as a structure, with headings and sub-headings.			
The browser search function works very well			
The browser has everything I need to search, navigate and display the Internet.			

Please comment on your ratings:

Usability: The ability to be used easily. *Similar terms*: ease of use, simplicity, parsimony, efficiency, user friendliness.	1= Strongly Disagree, 2= Disagree, 3= In the middle, 4=Agree, 5= Strongly Agree		
	Is the statement		
Browser Statements	**Clear?**	**Valid?** by the definition	**Important?** in browser selection
I accomplish my tasks easier and quicker with this browser.			
I did not need training to use it the first time.			
The user interface is consistent and easy to learn.			
A user who has never seen the interface before can accomplish most tasks intuitively.			
If a user has used it once before, they can always remember how to use it again.			
It uses words and concepts familiar to users.			
It is easy to get it to do what I want it to do.			
Its user messages are always clear and understandable.			
It is easy to become skillful at using it.			
It "remembers" previous actions, so when I type in something it often successfully completes it for me.			
It remembers my passwords and automatically enters them when a web site asks for them.			
The toolbar buttons are large and clear, with both text and a meaningful graphic			
The interface looks and feels nice.			
The interface is simple and not cluttered.			
It can auto-fill forms with common data like my name and address, so I don't have to keep typing it in.			
The help system is very good.			

Please comment on your ratings:

APPENDIX B

PREFERENCE QUESTIONNAIRE

This appendix contains the questionnaires used for investigating various aspects of user evaluation method preference in Research Question 2.

EVALUATION METHOD PREFERENCE QUESTIONNAIRE

1. Would you prefer to choose the criteria to evaluate with from a) all eight WOSP criteria, or would you prefer to regularly choose from b) a set of fewer criteria?

a) Choose from all eight
b) Choose from fewer – if so, specify how many 1-2-3-4-5-6-7 by ticking the appropriate box, e.g. if you would prefer to regularly evaluate using a set of two criteria, tick 2

7	6	5	4	3	2	1

Please briefly explain your response below:
....

2. If fewer than 8, tick which ones you would prefer to use:

Flexibility	Reliability	Security	Extendibility	Privacy	Functionality	Connectivity	Usability

Please briefly explain your response below:
......

3. Suppose you only used the criteria Functionality and Usability to select software. How would this compare in the following ways with using the eight criteria WOSP to select software?

1) ***Confidence,*** *that the method would give the best software choice.*
   a) For the two criteria Functionality/Usability evaluation I would be:

Very confident (5)	Confident (4)	Neutral (3)	Unsure (2)	Very unsure (1)

   b) For the eight criteria WOSP evaluation I would be:

Very confident (5)	Confident (4)	Neutral (3)	Unsure (2)	Very unsure (1)

2) ***Accuracy**, that the method correctly measured software performance.*
   a) I would feel the two criteria Functionality/Usability evaluation method was:

Very confident (5)	Confident (4)	Neutral (3)	Unsure (2)	Very unsure (1)

   b) I would feel the eight criteria WOSP evaluation was:

Very accurate (5)	accurate (4)	Neutral (3)	Inaccurate (2)	Very inaccurate (1)

3) ***Completeness**, that the method covered all the relevant factors.*
   a) The two criteria Functionality and Usability have all the dimensions I need to evaluate software:

Strongly agree (5)	agree (4)	Neutral (3)	Disagree (2)	Strongly disagree (1)

   b) The eight WOSP criteria have all the dimensions I need to evaluate software:

Strongly agree (5)	agree (4)	Neutral (3)	Disagree (2)	Strongly disagree (1)

Please briefly comment on your responses:

4. Supposing you were evaluating may different types of software. For each of the following criteria, state how often you would expect it to be a critical evaluation factor.

Criteria	***Scale (put an 'X' in the column that applies)***				
	Always or nearly always	*Very often*	*Often*	*Not often*	*Never or hardly ever*
Usability: *The ability to be used easily.* *Similar terms*: ease of use, simplicity, parsimony, efficiency, user friendliness.					
Functionality: *The ability to do the required functions* *Similar terms*: Effectiveness, capability, usefulness, power, strength.					
Connectivity: *The ability to exchange information with other systems.* *Similar terms*: interactivity, sociability, connectedness, networking, communication.					
Privacy: *The ability to limit unwanted information disclosure.* *Similar terms*: confidentiality, secrecy, camouflage, stealth, opaqueness.					
Extendibility: *The ability to make use of third party programs and data.* *Similar terms*: extendibility, compatibility, scalability, open standards, tool use.					
Security: *The ability to defend against hostile attack, including unauthorized entry, change, damage, hurt or takeover.* *Similar terms*: defendability, protectiveness, resistance to attack, resistance					
Reliability: *The ability to continue working despite errors or problems, or to quickly recover if failure occurs.* *Similar terms*: stability, dependability, recovery, durability, ruggedness					
Flexibility: *The ability to change itself to fit different situations.* *Similar terms*: adaptability, portability, customizability, plasticity, agility					

Please comment briefly on your response:

APPENDIX C

ORTHOGONAL STIMULI SET INCLUDING HOLDOUTS GENERATED USING SPSS

This appendix contains the SPSS-generated orthogonal stimuli set that was used to investigate the importance that users attach to the various WOSP criteria using the conjoint analysis method.

ORTHOGONAL STIMULI SET INCLUDING HOLDOUTS GENERATED USING SPSS

Browser No.	Privacy	Extendibility	Usability	Flexibility	Connectivity	Security	Functionality	Reliability
1	Medium	Low	Low	Medium	High	High	Medium	Low
2	Low	High	High	Medium	High	Medium	Low	High
3	Low	Low	High	Low	Medium	Medium	High	Medium
4	Medium	High	Medium	High	High	Low	Low	Medium
5	Low	Medium	Medium	High	Medium	High	Low	Medium
6	High	High	Low	Low	High	High	High	Low
7	Medium	Medium	High	Low	High	High	Low	Medium
8	High	Low	Low	High	Medium	High	Medium	High
9	High	Medium	High	High	Medium	Medium	High	Medium
10	High	Low	High	High	High	Low	Low	Low
11	High	Low	Medium	High	Low	Medium	High	Medium
12	High	Low	Medium	Low	Low	Medium	Low	High
13	Low	Medium	Low	High	High	Low	High	High
14	Medium	Low	High	Medium	Low	High	Medium	High
15	Medium	High	Medium	High	High	Low	Medium	Medium
16	High	High	Medium	Low	Medium	Medium	Low	High
17	Low	Medium	High	High	Low	Medium	Medium	Low
18	Low	Low	Medium	Low	High	High	Medium	High
19	High	Low	Low	High	High	Low	Medium	Medium
20	Medium	Medium	Medium	Low	Low	Low	High	High
21	Low	High	Medium	Medium	Low	High	High	Low

ORTHOGONAL STIMULI SET INCLUDING HOLDOUTS GENERATED USING SPSS (CONTINUED)

Browser No.	Privacy	Extendibility	Usability	Flexibility	Connectivity	Security	Functionality	Reliability
22	Medium	High	Low	High	Low	Medium	Low	High
23	Medium	Medium	Low	Low	Medium	Medium	Medium	Low
24	High	High	High	Low	Low	Low	Medium	Medium
25	High	Medium	Low	Medium	Low	High	Low	Medium
26	Medium	High	High	High	Medium	High	High	Low
27	High	Low	High	Medium	High	High	High	Low
28	Low	Low	Low	Low	Low	Low	Low	Low
29	High	Medium	High	Medium	Medium	Low	High	High
30	Low	High	Low	Medium	Medium	Low	Medium	Medium
31	High	Medium	Medium	Medium	High	Medium	Medium	Low
32	Medium	Low	Medium	Medium	Medium	Low	Low	Low
33	Medium	Low	Low	Medium	High	Medium	High	Medium

APPENDIX D

INSTRUMENT FOR DATA COLLECTION OF CONJOINT ANALYSIS

This appendix contains the instrument that was used to collect data from subjects for conjoint analysis in Research Question 3.

Subject ID: ________

RE: WEB BROWSER GRADING

Dear Senior Manager,
This is the last step to identify a suitable browser for our company from the many different types vendors offer. As you know, we recently asked middle managers to evaluate browser performance statements, and from that data selected the six most important for each of the eight factors of ***openness, security, connectivity, usability, privacy, flexibility, effectiveness,*** and ***reliability***.

This took a bit of time, as the evaluation from all the managers had to be received before the selection could be done. The evaluations given by the managers, and their comments were taken very seriously. Selection of the most important statements involved a lot of tradeoffs and compromises as to what the individual managers considered most important, in order to come up the statements that would be generally agreeable to all for the assessment of the performance factors.

The Procurement Department then asked an IT consultant to rate these factors as HIGH, MEDIUM, or LOW for each of 33 browsers and browser versions. The results are given below. For your convenience, the eight factors, each with the six main statements listed under it, are given in the next page, followed by the 33 browsers, given one per page. For each browser, the factors (and their descriptions) are given, and the consultant's rating of each factor for the browser.

The executive must now decide which browser to buy based on a rational process, so all senior managers have been asked to help choose. **Your input as a senior manager in this exercise is very important.** I ask you to use your personal judgment, as follows:

A. *Grade* each browser **Strong, Good, Adequate, Limited or Weak** based on its ratings. There are 33 browsers, each with a separate ratings page. Type in your final grade for each browser at the bottom of its page. The browsers are not listed in any particular order.

B. *Complete the summary table at the end* – carefully copy forward your grades, then give each browser a score from 1-100, and finally rank them from best to worst.

C. Please note that, although not anticipated, should there be a discrepancy between the evaluation of the factors as listed in the summary and that given for the individual browser pages, the summary's evaluation should prevail, and you should do the necessary correction on the browser page.

D. *Explain* why you decided as you didOnce again, thank you for your cooperation in this task.

Sincerely,

Niels Borg

CEO, QUANTUM COMPUTING INC, "Riding the e-light wave".

Please print this page and use it for reference while you are grading the browsers.

Connectivity	Functionality	Flexibility	Extendibility
It gives access to other ways of communicating, like Telnet, Ftp, email and chat.	The application's search function works very well	It runs on all our computers and operating systems.	It follows all World Wide Web source code and data standards, e.g. unicode.
If importing a file is stopped, or fails for any reason, it can be restarted again later from where it left off, saving a lot of time.	I can do what I want to do quickly with this application.	I can change the toolbar to include any buttons in any order.	It can install plug-ins for the latest application components and modules.
When importing files, it gives useful information, like the estimated time to complete the importing of a large file.	Overall, a system's functionality affects its performance.	It can be used in any organizational environment without much modification.	It can save files to my computer in most of the other popular formats.
Overall, a system's connectivity affects its performance.	Using it in my work increases my productivity.	Its settings can be changed easily and quickly.	It works with all popular third party tools.
It tells me if new software version upgrades are available to download.	The application improves my work performance.	It is easily changed to fit disability needs e.g. larger text or graphics for those with poor vision.	It can handle graphics, sound and video in a wide variety of different formats.
When importing a very "rich" web site, it tells you what percentage has loaded, and how long to go	I can work on many tasks at once, and jump between them by clicking on a tab or button.	Overall, a system's flexibility affects its performance.	I can clip and paste from the application into any company software application.

Reference page for grading browsers (continued)

Privacy	Reliability	Security	Usability
Password information always shows as asterisks, so others cannot look over my shoulder to see them.	It never breaks down or "hangs" (fails to respond), even if I use it for a long time.	It can protect itself from outside attacks, or things that take over control, like pop-up ads.	I accomplish my tasks easier and quicker with this application.
Any sensitive information I give the application, like logon passwords, is encrypted, so others can't see it.	Even if I multi-task, and do many things at once, it still works well.	Its security features are easy to use and are effective.	The user interface is consistent and easy to learn.
It tells me if my computer is being used to record information about me or my computer.	Overall, a system's reliability affects its performance.	It tells me if a file attempts to change any of my settings, and can prevent it.	It is easy to become skillful at using it.
It tells me clearly what is being stored on my hard drive, e.g. in cookies.	It can backup its data to a single file that can be reloaded if there is a problem	It can detect and prevent spyware from installing.	A user who has never seen the interface before can accomplish most tasks intuitively.
It prevents online tracking, spam harvesting, and snoops while I am connected to the Internet.	The number of errors while using it is low, or lower than average.	When a file is downloaded to the hard drive, it is checked for viruses before use.	The interface is simple and not cluttered.
It gives an option for removing personal information from a file on saving.	I am never stuck, because there is always another way to do the same thing, e.g. mouse click or keyboard.	It gives me a choice to allow extensions like java applets or ActiveX programs to run or not.	Overall, a system's usability affects its performance.

INDIVIDUAL GRADING: BROWSER 1

Connectivity	Privacy	Reliability	Extendibility
The ability to manage channels of communication to exchange meaning and information. *Similar terms*: interactivity, sociability, connectedness.	**The ability to prevent unwanted information disclosure.** *Similar terms*: confidentiality, secrecy.	**The ability to continue to work, despite errors, failures, or load.** *Similar terms*: stability, dependability	**The ability to make use of third party programs and data.** *Similar terms*: openness, compatibility, scalability, standardized.
High	*High*	*Medium*	*Low*
Functionality	**Usability**	**Security**	**Flexibility**
The ability to perform the required functions. *Similar terms*: effectiveness, capability, usefulness.	**The ability to be used easily.** *Similar terms*: ease of use, simplicity, parsimony, efficiency.	**The ability to defend against hostile attack, including unauthorized entry, change, damage, hurt or takeover.** *Similar terms*: defendability, protectiveness	**The ability to change to fit different situations.** *Similar terms*: adaptability, portability, customizability, plasticity
Medium	*Low*	*Low*	*High*

Based on the above ratings, I grade this browser:

………

(Type Strong, Good, Adequate, Limited or Weak where the dots are)

B. BROWSER GRADING SUMMARY TABLE

a. *Copy your browser grades forward* from the previous pages into the Grade column, as **S, G, A, L or W** (for Strong, Good, Adequate, Limited or Weak).

b. *Score the browsers 1 – 100 in the Score column.* Pick what you think is the worst browser and score it 1. Then pick the best browser and give it 100. **Now give the other browsers scores in between 1 and 100.** You may use your browser grade as a guide, e.g.

   i. *Weak*, score from 1- 20.

   ii. *Limited*, score from 21- 40

   iii. *Adequate*, score from 41- 60.

   iv. *Good*, score from 61- 80

   v. *Strong*, score from 81- 100.
   However you must spread the scores from 1 to 100. So if you did not grade any browser as "Strong", you might score the "Good" browsers from 61-100. The important point is to score all the browsers from 1 to 100, the weakest being 1, and the strongest 100. Also, these scores need not be evenly distributed among the grades, or even within a grade.

Rank the browsers from 1 to 33 in the Rank column, with 1 being the best and 33 the worst. No two browsers should have the same rank.

No.	**Privacy**	**Extendibility**	**Usability**	**Flexibility**	**Connectivity**	**Security**	**Functionality**	**Reliability**	**Grade**	**Score 1-100**	**Rank**
1	High	High	Medium	Low	Medium	Low	Low	High			
2	High	Medium	Low	Low	Medium	Low	High	Medium			
3	High	High	Low	Low	Low	High	Low	High			
4	High	High	Low	Medium	Medium	Medium	Medium	Medium			
5	Medium	Medium	Low	High	High	High	High	High			
6	Low	Low	Low	Low	Low	Low	Low	Low			
7	Low	Low	Low	Medium	Medium	High	Medium	High			
8	High	Medium	Medium	Medium	Low	High	High	Low			

No.	Privacy	Extendibility	Usability	Flexibility	Connectivity	Security	Functionality	Reliability	Grade	Score 1-100	Rank
9	Low	Medium	High	High	Low	Low	Medium	High			
10	Medium	High	High	Medium	High	High	Low	Medium			
11	Low	High	High	Low	Low	Medium	Medium	Low			
12	Medium	Low	Medium	Medium	Low	Medium	High	High			
13	Medium	High	High	High	Low	Medium	Medium	Low			
14	Medium	High	High	Low	Medium	Low	High	High			
15	High	Medium	Medium	High	Low	Medium	Low	High			
16	Medium	Medium	Low	Low	Low	Medium	Low	Medium			
17	Low	High	Medium	Medium	Low	Low	High	Medium			
18	Medium	High	Medium	Medium	High	High	Medium	High			
19	High	Medium	Medium	High	Medium	Medium	Low	High			
20	High	High	Low	Low	High	High	High	Medium			
21	High	Low	High	High	Low	High	Medium	Medium			
22	Medium	Medium	Low	Medium	Medium	Low	Medium	Low			
23	Low	High	Medium	High	Medium	High	Low	Low			
24	High	Low	High	Medium	High	Low	Low	High			
25	Low	High	Medium	Low	High	Medium	Medium	High			
26	Low	Medium	High	Low	Medium	High	High	Medium			
27	High	Low	High	Low	Medium	Medium	High	Low			
28	Medium	Low	Medium	Low	High	High	Medium	Low			
29	High	High	Low	High	High	Low	High	Low			
30	Low	Low	Low	High	High	Medium	High	Medium			
31	High	Medium	Medium	Low	High	Low	Medium	Medium			
32	Medium	Low	Medium	High	Medium	Low	Low	Medium			
33	Low	Medium	High	Medium	High	Medium	Low	Low			

C. **EXPLAIN**: Write a paragraph to explain how you made your selection...

REFERENCES

Agarwal, R. and Venkatesh V. (2002). Assessing a firm's web presence: A heuristic evaluation procedure for the measurement. *Information Systems Research,* 13(2), 168-186.

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes,* 50(2), 179-211.

Applegate, L. M., McFarlan, F.W., and McKenney, J.L. (1996). Corporate Information Systems. Boston.

Araujo, I. and Araujo, I. (2003). Developing trust in internet commerce. *Proceedings of the 2003 Conference of the Centre for Advanced Studies on Collaborative Research,* 1-15

Arlitt, M., Krishnamurthy, D. and Rolia, R. (2001). Characterizing the scalability of a large web-based shopping system. *ACM Transactions on Internet Technology (TOIT),* 1(1), 44-69.

Avizienis, A., Laprie, J.C. Randall, B. and Landwehr, C. (2004). Basic concepts and taxonomy of dependable and secure computing. *IEEE Transactions on Dependable Computing* 1(1), 11-33.

Ba, S., Stallaert, J., and Whinston, A.B. (2001). Research commentary: Introducing a third dimension in information systems design - the case for incentive alignment. *Information Systems Research,* 1(3), 225-239

Barnett, R. C., and Marshall, N. L.(1981) The relationship between women's work and family roles and their subjective well-being and psychological distress, *Women, Work and Health: Stress and Opportunities,* M. Frankenhaeuser, V. Lundberg, and M. A. Chesney (eds.), Plenum, New York, 111-136

Bajaj, A. (2000). A Study of Senior Information Systems Managers' Decision Models in Adopting New Computing Architectures. *Journal of the Association of Information Systems (JAIS),* 1(4).

Bell, D. and LaPadula, L. (1973). Secure computer systems: mathematical foundations and model. *MITRE Report M74-244*, MTR 2547 2.

Bem, D. J. and Allen, A. (1974). On predicting some of the people some of the time: the search for cross-situational consistencies in behavior, *Psychological Review* (81:6), 506-520.

Bergeron, F., Rivard, S., and De Serre, L. (1990). Investigating the support role of the information center, *MIS Quarterly* (14:3). 247-259. Bertalanffy, L. V. (1968). General System Theory. New York: George Braziller Inc.

Beyers, W. B. and Lindahl, D. P. (1999). Workplace flexibilities in the producer services. *The Service Industries Journal*, 19(1), 35-60.

Bidgoli, H. (1996). A new productivity tool for the 90's: group support systems. *Journal of System Management,* 47(4), 56-62.

Blignaut, P. J. (2004). Computerized self-administered questionnaires on touchscreen kiosks: do they tell the truth? *2004 conference on Human factors and computing systems,* 1183-1186

Boehm, B. (1981). Software Engineering Economics. Englewood Cliffs, NJ.

Bondi, A. B. (2000). Characteristics of scalability and their impact on performance. *Proceedings of the Second International Workshop on Software and Performance,* 195-203.

Borning, A. (1987). Computer system reliability and nuclear war. *Communications of the ACM*, 30(2), 112-131.

Bowen, W. (1986). The puny payoff from office computers. *Computers in The Human Context: Information, Productivity, and People,* 267-271.

Bozionelos, N. (1996). Psychology of computer use: prevalence of computer anxiety in british managers and professionals. *Psychological Reports* (78:3), 995-1002.

Brocklehurst S., Kanoun K., Laprie J.C., Littlewood B., Metge S., Mellor P. et al. (1991). Reliability analyses of workstation failure data. *Rapport LAAS No91172 ESPRIT'91 Conference,* Brussels, 806-821.

Browne, G. J. and M. B. Rogich M. B. (2001). An empirical investigation of user requirements elicitation: comparing the effectiveness of prompting techniques. *MIS Quarterly*, 17(4).

Brynjolfsson, E. (1993). The productivity paradox of information technology. *Communications of the ACM* 36(12), 66-77.

Brynjolfsson, E. (1998). Beyond the productivity paradox. *Communications of the ACM*, 41(8), 49-55.

Butler, A., S. (2002). Software evaluation: security attributes evaluation method: a cost-benefit approach. *Proceedings of the 24th International Conference on Software Engineering,* 232-240.

Byrd, T. A. and D. E. Turner (2000). Measuring the flexibility of information technology infrastructure: exploratory analysis of a construct. *Journal of Management Information Systems*, 17(1), 167-208.

Chakrabarti, A. K. (1974). The role of champion in product innovation. *California Management Review*, 17(2),58-62.

Churchman, C. W. (1979). *The Systems Approach.* New York, Dell Publishing.

Clarke, C. L. A., Tilker, P.L., Tran, A. Q-L, Harris, K., and Cheng, A. S. (2003). A reliable storage management layer for distributed information retrieval systems. *ACM Proceedings of the Twelfth International Conference on Information and Knowledge Management,* 207-215.

Clarke, R. (1999). Internet privacy concerns confirm the case for intervention. *Communications of the ACM*, 42(2), 60-67.

Compeau, D. R. and Higgins, C. A. (1995). Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly*, 19(2), 189-211.

Cortellessa, V., Singh, H., and Cukic, B. (2002). Early reliability assessment of UML based software models. *Proceedings of the third international workshop on Software and performance,* 302-309.

Curley, K. F. (1984). Are there any real benefits from office automation? *Business Horizons,* 27(4), 37-42.

Davis, A. (1993). *Software Requirements Objects, Functions, and States.* Englewood Cliffs, NJ.

Davis, F. D. (1986). A technology acceptance model for empirically testing new end-user information systems: theory and results. *MIT Sloan School of Management. Cambridge, MA, Massachusetts Institute of Technology.*

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly,* 13(3.),319-339.

Davis, F. D., Bagozzi, R. P., and Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. *Management Science*, 35 (8), 982-1004.

Davis, F. D., Bagozzi, R.P., and Warshaw, P.R. (1992). Extrinsic and intrinsic motivation to use computers in the workplace. *Journal of applied social psychology,* 22(14), 1111-1132.

Davis, S. (1987). *Future Perfect.* Reading MA.

De Simone, M. and Kazman, R. (1995). Software architectural analysis: an experience report. *Proceedings of the 1995 conference of the Centre for Advanced Studies on Collaborative Research,* 251-261.

Denaro, G., Polini,A. and Emmerich, W. (2004). Early performance testing of distributed software applications. *ACM SIGSOFT Software Engineering Notes, Proceedings of the fourth international workshop on Software and performance,* 94-103.

Dennis, A. R. and Valacich, J. S. (1999). Rethinking media richness: towards a theory of media synchronicity. *Proceedings of the 32nd Hawaii International Conference on System Sciences, Hawaii,* 1017.

Desanctis, G. and Poole, M. S. (1994). Capturing the complexity of advanced technology use: adaptive structuration theory. *Organizational Science,* 5(2),121-147.

Devito, J. A. (1997). *Human Communication.* New York, Longman.

Dourish, P. and Redmiles D. (2002). An approach to usable security based on event monitoring and visualization. *Proceedings of the 2002 workshop on New security paradigms.*

EC (1993). Information Technology Security Evaluation Criteria.

El Sawy, O. A. and Nanus, B. (1989). Toward the design of robust information systems. *Journal of Management Information Systems,* 5(4), 33-54.

Eloff, J. and Eloff, M. (2003). Information security management: a new paradigm. *Proceedings of the 2003 annual research conference of the South African institute of computer scientists and information technologists on Enablement through technology,* 130-136.

Farahmand, F., Navathe, B.S.,Enslow, P.H., and Sharp, G.P. (2003). Managing vulnerabilities of information systems to security incidents. *Proceedings of the 5th international conference on Electronic commerce,* 348-354.

Farbey, B., Land F., and Targett, D. (1993). *How to Assess Your IT Investment.* Oxford, Butterworth Heinemann.

Fischoff, B. (1989). Eliciting knowledge for analytical representations. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(3), 448-461.

Fishbein, M. and Ajzen, I. (1975). *Belief, Attitude, Intention and Behavior: an Introduction to Theory and Research.* Reading, MA, Addison-Wesley.

Fule, P. and Roddick, J. (2004). Detecting privacy and ethical sensitivity in data mining results. *Proceedings of the 27th conference on Australasian computer science 26,* 159-166.

Garfinkel, S. L., Margrave,D., Schiller, J.I., Nordlander, E., and Miller, R.C. (2005). How to make secure email easier to use. *Proceeding of the SIGCHI conference on Human factors in computing systems, Portland Oregon,* 701-710.

Gavison, R. (1984). *Privacy and the Limits of the Law. Philosophical Dimensions of Privacy: An Anthology.* F. Schoeman. Cambridge, Cambridge university press: 346-402.

Goan, T. (1999). A cop on the beat: collecting and appraising intrusion evidence. *Communications of the ACM,* 42(7), 62-69.

Goodhue, D. L. and Thompson, R.L. (1995). Task-technology fit and individual performance. *MIS Quarterly,* 19(2), 213-236.

Goodwin, N. C. (1987). Functionality and usability. *Computing Practices,* 30(3),229-233.

Gray, R. S. (2000). Soldiers, agents and wireless networks: a report on military application. *Proceedings of the 5th International Conference and Exhibition on the Practical Application of Intelligent Agents and Multi-Agents,* 1-6.

Gray, R. S., Kotz, D., Newport, C., Dubyovsky, N., Fiske, A, Liu, J.et al. (2004). Outdoor experimental comparison of four ad hoc routing algorithms. *Proceedings of the 7th ACM International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems,* 220-229.

Gray, W. D. and M. C. Salzman, M. C. (1998). Damaged merchandise? A review of experiments that compare usability evaluation methods. *Human Computer Interaction,* 13(3), 203-261.

Hall, D., and Mansfield, R. (1995). Relationships of age and seniority with career variables of engineers and scientists. *Journal of Applied Psychology* (60:2), 201-210.

Hamilton, J. S. (1981). *A Survey of Data Processing Post Installation Evaluation Practices.* Univ. of Minnesota.

Heinl, P., Horn, S., Jablonski, S., Neeb, J., Stein, K., and Teschke, M. (1999). A comprehensive approach to flexibility in workflow management systems. *ACM SIGSOFT Software Engineering Notes, Proceedings of the International Joint Conference on Work Activities Coordination and Collaboration,* 24. (2), 79-88.

Hix, D. and Schulman, R. S (1991). Human-computer interface development tools: a methodology for their evaluation. *Communications of the ACM,* 34, 74-87.

Hofstadter, D. R. (1999). *Gödel, Escher, Bach: An Eternal Golden Braid.* New York.

Hu, P. J., Chau, P.Y.K., Sheng, O.R.L, and Tam, K.Y. (1999). Examining the technology acceptance model using physician acceptance of telemedicine. *Journal of Management Information Systems,* 16(2),91-112.

Ian, M. (1989). Computing matters: making it work to best effect. *Management Today,* 109-110.

IBM Software Engineering Using Hyperspaces, *IBM Research, MDSOC.*

InformationWeek (1999). Bond the new and the old: Enterprise architecture. *InformationWeek I,* 108-109.

Irvine, C. and Levin, T. (1999). Towards a taxonomy and costing method for security services. *Proceedings of the 15th Annual Computer Security Applications Conference,* 183.

ISO (1989). ISO 7498-2:1989. *Information processing systems - Open Systems Interconnection, Basic Reference Model, Part 2: Security Architecture.*

Jajodia, S. (1996). Database security and privacy. *ACM Computing Surveys (CSUR),* 28(1), 129-131.

Janis, I. L. and Mann, L. (1977). *Decision Making.* New York, Free Press.

Jokela, T., Iivari,N., Matero, J. and Karukka, M. (2003). The standard of user-centered design and the standard definition of usability: analyzing ISO 13407 against ISO 9241-11. *Proceedings of the Latin American conference on Human-computer interaction,* 53-60.

Jonsson, E. (1998). An integrated framework for security and dependability. *Paper presented at the NSPW, Charlottsville, VA, USA.*

Joshi, J. B. D., Aref, W.G., Ghafoor, A., and Spafford, E. H. (2001). Security models for web-based applications. *Communications of the ACM*, 44(2), 38-44.

Kahneman, D., Slovic, P., and Tversky, A. (1982). *Judgment Under Uncertainty.* New York, Cambridge University Press.

Karat, J. (1997). Evolving the scope of user-centered design. *Communications of the ACM*, 40(7), 33-38.

Kelman, H. C. (1958). Compliance, identification and internalization: three processes of attitude change? *Journal of Conflict Resolution,* 2, 51-60.

Kim, K. and Ellis, A.E. (2001). Workflow performance and scalability analysis using the layered queuing modeling methodology. *Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work,* 135-143.

Kjaer, A. and Madsen, K.H. (1995). Participatory analysis of flexibility. *Communications of the ACM,* 38(5), 53-60.

Knoll, K. and Jarvenpaa, S. L. (1994). Information technology alignment or "fit" in highly turbulent environments: the concept of flexibility. *Proceedings of the 1994 Computer Personnel Research Conference on Reinventing IS: Managing Information Technology in Changing Organizations,* 1-14.

Kock, N. (2001). The ape that used e-mail: understanding e-communication behavior through evolution theory. *CAIS Journal,* 5 (3), 1-29.

Kraut, R., J. Galegher,J., Fish, R., and Chalfonte, B. (1992). Task requirements and media choice in collaborative writing. *Human Computer Interaction,* 7, 375-407.

Krishna, M. S. and Subramanyam, R. (2004). Quality dimensions in e-commerce software tools: an empirical analysis of North American and Japanese markets. *Journal of Organizational Computing and Electronic Commerce*, 14(4), 223-241.

Kumar, K. (1990). Post implementation evaluation of computer-based information systems: current practices. *Communications of the ACM*, 33(2), 203-212.

Lamb, R. and Kling. R. (2003). Reconceptualizing users as social actors in information systems research. *MIS Quarterly,* 27(2), 197-235.

Lamsweerde, A. V. (2000). Formal specification: a roadmap, FOSE 00. *International Conference on Software Engineering, ACM Press,* 149-159.

Laprie, J. C. (1992). *Dependability: basic concepts and terminology.* Springer-Verlag 5.

Law, D. R. (1998). Scalable means more than more: a unifying definition of simulation scalability. *Proceedings of the 30th Conference on Winter Simulation,* 781-788.

Lea, M. (1991). Rationalist assumptions in cross-media comparisons of computer-mediated communication. *Behavior and Information Technology,* 10(2),153-172.

Lecerof, A. and Paterno, F. (1998). Automatic support for usability evaluation. *IEEE Transactions on Software Engineering,* 24 (10), 863-888.

Lee, S. and Leifer, R.S. (1992). A framework for linking the structure of information systems with organizational requirements for information sharing. *Journal of Management Information Systems,* 8(4), 27- 45.

Leonard-Barton, D. (1988). Implementation characteristics of organizational innovations. *Journal of Communications Research*, 15(5), 603-631.

Leonard-Barton, and I. Deschamps, D. (1988). Managerial influence in the implementation of new technology. *Management Science*, 34(10), 1252-1265.

Levy, J. A.(1988). Intersections of gender and aging. *The Sociological Quarterly* (29:4), 479-486.

Lim, K. H. and Benbasat, K. (2000). The effect of multi-media on perceived equivocality and perceived usefulness of information systems. *MIS Quarterly,* 24(3), 449-471.

Littlewood, B. and Strigini, L. (2000). Software reliability and dependability: a roadmap. *Proceedings of the conference on The future of Software engineering,* 175-188.

Losavio, F., Chirinos,L., Matteo, A., Levy, N., Ramdane-cherif, A. (2004). Designing quality architecture: incorporating ISO standards into the unified process. *Information Systems Management*, 21(1), 27-44.

Lubinski, D., Tellegen, A. and Butcher, J. N. (1983) Masculinity, femininity, and androgyny viewed and assessed as distinct concepts. *Journal of Personality and Social Psychology* (44:2), 428-439.

Malhotra, Y. and Galletta, D.F. (1999). Extending the technology acceptance model to account for social influence: theoretical bases and empirical validation. *Proceedings of the 32nd Hawaii International Conference on System Sciences, University of Hawaii,* 1006.

Maturana, H. R. and Varela, F. J. (1998). *The Tree of Knowledge: The biological roots of human understanding.* Boston, MA, Shambala Publications.

Merriam-Webster (2002). *Webster's Third New international Dictionary Unabridged.* Springfield, MA, Merriam-Webster, Inc.

Meyer, B. (1987). Eiffel: programming for reusability and extendibility. *ACM SIGPLAN Notices,* 22 (2), 85-94.

Miller, J. B. (1976). *Toward a New Psychology of Women.* Beacon Press, Boston.

Minton, H. L., and Schneider, F. W. (1980) *Differential Psychology.* Waveland Press, Prospect Heights, IL.

Montz, A. B. and Peterson, L. (1998). Controlled flexibility in systems design. *Proceedings of the 8th ACM SIGOPS European Workshop on Support for Composing Distributed Applications,* 247-254.

Moore, G. C. and Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research,* 2 (3), 192-222.

Moreira, A., Araujo,J., and Brito, I. (2002). Crosscutting quality attributes for requirements engineering. *ACM,* 27,167-174.

Muftic, S. (1989). *Security Mechanisms for Computer Networks.* England, Ellis Horwood Ltd.

Neumann, P. G. (1995). *Computer Related Risks.* Reading, MA, Addison-Wesley.

O'Reilly, C. I. and Chatman, J. (1986). Organizational commitment and psychological attachment: the effects of compliance, identification, and internalization of prosocial behavior. *Journal of Applied Psychology,* 71 (3), 492-499.

Orlikowski, W. J. and Barley, S. R. (2001). Technology and institutions: what can research on information technology and research on organizations learn from each other? *MIS Quarterly,* 25, (2).

Perrow, C. (1984). *Normal Accidents: Living With High-Risk Technology. New York,NY,* Basic Books.

Plude, D., and Hoyer, W. (1985).Attention and performance: identifying and localizing age deficits. *Aging and Human Performance,* N. Charness (ed.), John Wiley & Sons, New York. 47-99.

Prayurachatuporn, S. and Luigi Benedicenti, L. (2001). Increasing the reliability of control systems with agent technology. *ACM SIGAPP Applied Computing Review,* 9 (2), 6-12.

Ramjee, R. and La Porta, T. L. (1998). Performance evaluation of connection rerouting schemes for ATM-based wireless networks. *IEEE/ACM Transactions on networking,* 6(3), 249-261.

Rana, O. F. and Stout, K. (2000). What is scalability in multi-agent systems? *Proceedings of the fourth international conference on Autonomous agents,* 56-63.

Rice, R. (1987). Computer-mediated communication and organizational innovation. *Journal of Communication,* 37 (4), 65-95.

Rice, R. (1994). *Network analysis and computer-mediated communication systems. Advances in Social Network Analysis.* S. Wasserman and J. Galaskiewicz. Newbury Park, CA, Sage.

Rinard, M. (2003). Acceptability-oriented computing. *ACM SIGPLAN Notices,* 38 (12), 57-75.

Rivard, S., Poirier, G., Raymond, L., and Bergeron, F. (1997). Development of a measure to assess the quality of user-developed applications. *ACM SIGMIS Database,* 28 (3), 44-58.

Robinson, W. N., Pawloski, S.D. and Volkov, V. (2000). Requirements interaction management. *ACM Computing Surveys,* 35 (2), 132-190.

Rogers, E. (1995). *Diffusion of Innovations.* New York, Free Press.

Rombel, A. (2001). Privacy and security in a wired world. *Global Finance,* 15 (1), 28.

Seddon, P. B., Graeser,V., and Willcocks, L. P. (2002). Measuring organizational IS effectiveness an overview and update of senior management perspectives. *ACM SIGMIS Database,* 33 (2), 11-28.

Sethi, V. and King, W. R. (1994). Development measures to assess the extent to which an information technology application provides competitive advantage. *Management Science,* 40 (12), 1601-1627.

Shah, S. H., Chen, K., and Nahrstedt, K. (2005). Dynamic bandwidth management for single-hop ad hoc wireless networks. *Communications of the ACM,* 10(1-2): 199-217.

Sharda, R., Barr, S.H., and McDonnell J. (1988). Decision support system effectiveness: a review and empirical test. *Management Science,* 34 (2), 139-159.

Sheppard, B. H., Hartwick, J. and Warshaw, P. R. (1988). The theory of reasoned action: a meta-analysis of past research with recommendations for modifications and future research. *Journal of Consumer Research,* 15 (3), 325-343.

Shneiderman, B. (2000). Universal usability. *Communications of the ACM,* 43 (5), 84-91.

Shneiderman, B. (2003). Promoting universal usability with multi-layer interface design. *Proceedings of the 2003 Conference on Universal Usability,* 1-8.

Simon, H. (1997). *Administrative Behavior.* New York, Free Press.

Singh, S. and Kotzé, P. (2002). Towards a framework for e-commerce usability. *Proceedings of the 2002 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on Enablement through Technology,* 2-10.

Sommerville, I. (2000). *Software Engineering.* Pearson Education.

Sommerville, I. (2004). *Software Engineering.* Upper Saddle River NJ, Pearson Education.

Standish (1996). *Chaos.* The Standish Group International, Inc.

Sylla, C. and Wen, H. J. (2002). A conceptual framework for evaluation of information technology investments. *Int. J. Technology Management*, 24(2), 236-261.

Tajfel, H. (1978). *Differentiation between Social Groups: Studies in the Social Psychology of Intergroup Relations.* London. Academic Press.

Taylor, S. and Todd, P. A. (1995). Understanding information technology usage: a test of competing models. *Information Systems Research,* 6, (2), 144-176.

Thompson, R. L., Higgins, C.A., and Howell, J.M. (1991). Personal computing: Towards a conceptual model of utilization. *MIS Quarterly,* 15, (1), 125-142.

Torrellas, G. A. S. and Vargas,L. A. V. (2003). Modelling a flexible network security system using multi-agents systems: security assessment considerations. *Proceedings of the 1st International Symposium on Information and Communication Technologies,* 365-371.

Triandis, H. C. (1977). *Interpersonal Behavior.* Monterey, CA.

Turban, E. (1995). *Decision Support Systems and Expert Systems.* Prentice Hall.

Turban, E., Lee, J.K., and Chung, M. (2000). *Electronic Commerce: Managerial Perspective.* Upper Saddle River, NJ, Prentice Hall.

Vallerand, R. J. (1997). *Toward a hierarchical model of intrinsic and extrinsic motivation. Advances in Experimental Social Psychology.* M. Zanna. New York, Academic Press. 29, 271-360.

Venkatesh, V.,Morris, M. G., Davis, G.B., and Davis, F. D. (2003). User acceptance of information technology: toward a unified view. *MIS Quarterly,* 27 (3), 425-478.

Venkatesh, V. and C. Speier, C. (1999). Computer technology training in the workplace: a longitudinal investigation of the effect of the mood. *Organizational Behavior and Human Decision Processes,* 79 (1), 1-28.

Verykios, V. S., Bertino, S. E., Fovino, I.N., Provenza, L.P., Saygin, Y., and Theodoridis, Y. et al. (2004). State-of-the-art in privacy preserving data mining. *ACM Computing Surveys (CSUR),* 33 (1), 50-57.

Vessey, I. and Galletta, D. F. (1991). Cognitive fit: an empirical study of information acquisition. *Information Systems Research,* 2 (1),63-84.

Vickery, S., Calantone, R., and Droge, C. (1999). Supply chain flexibility: an empirical study. *Journal of Supply Chain Management,* 35 (3), 16-24.

Vining, G.G. (1998). *Statistical Methods for Engineers.* Pacific Grove, CA, Duxbury Press

Walther, G. H. and O'Neil, H.F.J. (1974). Online user-computer interface - the effects of interface flexibility, terminal type, and experience on performance. *AFIPS National Computer Conference Proceedings,* 379-384.

Wang, H., Lee, M. K.O., and Wang, C. (1998). Consumer privacy concerns about Internet marketing. *Communications of the ACM,* 41 (3), 63-70.

Ware, W. H. (1984). Information systems security and privacy. *Communications of the ACM,* 27 (4), 315-321.

Wendelken, S., McGrath, S. and Blike, G. T. (2003). A medical assessment algorithm for automated remote triage. *Proceedings of the 25th annual engineering in medicine and biology conference, Cancun, Mexico.*

Westland, J. C. and Clark, T.H.K. (2000). *Global Electronic Commerce.* Cambridge, MA, MIT Press.

Whittaker, J. A. and Voa, J. (2000). Towards a more reliable theory of software reliability. *IEEE Computer Society*, 36-42.

Whitworth, B., Gallupe, B., and McQueen, R. (2000). A cognitive three process model of computer-mediated groups: theoretical foundations for groupware design. *Group Decision and Negotiation,* 9 (5), 431-456.

Whitworth, B. and M. Zaic (2003). The WOSP model: balanced information system design and evaluation. *Communications for the Association for Information Systems,* 12, 258-282.

Woods, D. D. (1993). The price of flexibility. *Proceedings of the 1st International Conference on Intelligent User Interfaces,* 19-25.

Wright, M. A. and Kakalik, J.S. (1997). The erosion of privacy. *ACM SIGCAS Computers and Society,* 27 (4), 22-25.

Young, T. R. (1984). The lonely micro. *Datamation,* 30 (4), 100-114.