EXAMINATION OF THE RELATIONSHIP BETWEEN ATTITUDE AND PERCEIVED USABILITY OF SOFTWARE IN A CONTEXT-DRIVEN SPECIFIC SCENARIO

ALP YÖRÜK

M.B.A., Business Administration, Işık University 2002

B.S., Mechanical Engineering, Istanbul Technical University, 1999

Submitted to the Graduate School of Social Sciences

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Contemporary Management Studies

IŞIK UNIVERSITY

2008

EXAMINATION OF THE RELATIONSHIP BETWEEN ATTITUDE AND PERCEIVED USABILITY OF SOFTWARE IN A CONTEXT-DRIVEN SPECIFIC SCENARIO

Abstract

Literature regarding human and computer interaction is enormous in size. Since computers have started to play an important role in our everyday life, researchers have focused on improving the interaction between these two entities. The usability of computer & software systems has been the main area of interest. Various techniques have been developed to measure usability. Some researchers have focused on objective usability whereas some have focused on the psychological aspect by examining subjective usability. When the objective is to examine a potential set of users' perceived usability, the context dependent nature of this specific literature even added to the difficulties. This research supplies an in-depth view of usability and computer attitude literatures, and offers computer attitude as an important variable which impacts perceived usability. A field study has been conducted in a population of 26 universities' students in Istanbul to test the model. The research findings indicate that computer attitude is an important variable that does have a significant effect on perceived usability.

KULLANICININ BİLGİSAYAR TUTUMU İLE ALGILANAN KULLANICI DOSTLUĞU ARASINDAKİ İLİŞKİNİN SPESİFİK BİR SENARYO ÜZERİNDEN İNCELENMESİ

Özet

İnsan ve bilgisayar arasındaki etkileşimi inceleyen literature oldukça geniştir. Bilgisayarların günlük hayatımız içerisinde önemli bir rol oynamaya başladığı andan beri, araştırmacılar bu iki varlık arasındaki etkileşimi iyileştirme konusunda çalışmalara odaklanmışlardır. Bilgisayar & yazılım sistemlerinin kullanıcı dostluğu bu çalışmalara konu olmuş önemli odak noktalarından biridir. Kullanıcı dostluğu, gerek objektif, gerekse, subjektif; psikolojik boyutu ile incelenebilir. Amacın potansiyel bir kullanıcı kitlesinin algıladığı kullanıcı dostluğunun incelenmesi olduğu durumlarda, bu literatürün spesifik durumlar bazındaki doğası işin zorluk seviyesini de arttırır. Bu araştırma, kullanıcı dostluğu ve bilgisayar tutumu literatürlerinde geniş bir taramayı içermekte ve bilgisayar tutumunu, algılanan kullanıcı dostluğunu etkileyen önemli bir değişken olarak önermektedir. İstanbul'daki 26 üniversiteyi kapsayan ana kütleyi temel alan bir saha çalışması ile önerilen model test edilmiştir. Araştırma bulguları; bilgisayar tutumunun algılanan kullanıcı dostluğu üzerinde anlamlı bir etkisi olduğuna işaret etmektedir.

Acknowledgements

There are many people I would like to thank for their precious value and support throughout this entire journey. First of all, I thank Murat Ferman, my professor and dissertation supervisor. He has inspired me throughout this thesis both intellectually and motivationally. Without his precious support it would not have been possible to complete this work.

Many thanks to Metin Çakıcı, my professor in Master of Business Administration program. He has shared his expertise on marketing research and statistics and paved the way for me to materialize research model.

I would also like to thank to Emrah Cengiz, my professor in MBA program. He has supplied his objective and valuable views on the subject and made it possible to refine my approach to the research.

Above all, most valuable and precious, my mother, Nilgün Sarıcan… I owe everything, my entire being, my whole life to this wonderful lady. She has been there when there was nobody. She always has been an inspiration for anybody who has been around her.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Literature regarding computers is enormous in size. Since man has discovered this machine which is capable of making computations in an ever increasing performance level, computers and human interaction has been an area of interest. It has been researched by both technical and non-technical disciplines. Computer engineers, IT professionals, MIS graduates, psychologists, management gurus all have published numerous books and articles regarding this machine-to-man interaction. It has been a huge step for man to include computer in his every day activities. This machine which is composed of several electronic equipments has been a tool which was a high-performance calculator first and then throughout its life stage evolved its existence into even a decision making tool! When computers became every day tools of not only a couple of laboratories but of businesses and home users, the quality of interaction between these two entities became a great area of research. Since then man's ability to use this tool to full potential and being able to improve both objective performance and subjective perception of quality in use has been the main topic.

Usability of systems (both hardware and software) was and is a great area of research. Everyday companies throughout the globe are investing significant amounts of their resources to improve their products so that customers can find them more satisfactory. This matter is that much important that although a product can be performing satisfactorily on objective measures if the perception of quality of interaction by end-users is low, it does not succeed in markets. Therefore these companies, offering these specific services are trying their best to identify points of improvements which will make it possible for them to improve the end-users perceptions about their products & services.

Usability literature regarding computers has been thoroughly studied starting from the 1980s. Still today usability for a system is a literature which is of great debates. Even on a definition basis, there are several highly acclaimed definitions by respectable authors but still no consensus. It is the very nature of usability that makes it so difficult to come up with a single definition of the concept. The context[*] dependence of the concept adds to the difficulties to overcome these debates. Even usability and its meaning by different categories of users who are interacting with the very same system can be completely different. Such as an end-user is much more oriented with the interface-quality aspect where as a technical admin, who is responsible for daily maintenance of system equipment, is much more oriented with the ease of maintainability of the system.

When one is to focus on attitudes towards computers, the whole idea and its interpretation becomes even more complex. Attitudes can greatly affect one's behaviors towards an object. When this object is a machine which tries to "interfere" with your every day routine and way of life, the issue becomes even more important. Therefore computer attitude sometimes becomes as the main barrier which acts as the main supplement to resistance to change, both for home and also for business users.

This thesis mainly focuses on the relation between computer attitude and usability. Our objective is to explore both of these concepts and their related literature. However defining usability and creating a scenario which makes it possible to analyze this important relationship is a great challenge which also promises important findings. If the literature review makes it possible to analyze this

[*] The user, the environment, the system and every other variable that interacts with the system in interest.

relationship, one then can become curious to attitude and its impact upon perceived usability (subjective usability).

The author is a professional within the IT software industry. His main responsibilities are system analysis & design. He also develops applications with in a development team for gaining further knowledge regarding coding which makes it possible for him to improve his perception of developers and their tasks. By this way, he can design systems which are more positively welcomed by developers working in the project. His experience is on web-based solutions.

The teams that have worked with the author has developed several acclaimed web-based solutions including CRM systems, Business Flow Management systems, Document Management systems, Internet Sites, E-Commerce sites and etc,. Amongst their client lists, Turkey's top associations and companies are existent such as Vestel, ODD (Turkish Automotive Distributors Association) (imported automobiles are making up 70% of the market with ODD being the most important association), HES Cable (World's #3 cable manufacturer), Roman (One of the most famous woman-dress designer and chain stores owner in Turkey), Remax-Turkey, GEM (Global Equity Management, a Switzerland based equity management company that has big investments in pharmaceutical industry), DDF (One of the most famous advertising agency in Turkey), Fire of Anatolia ("Anadolu Ateşi", world wide known Turkish dance group) and several other companies.

Throughout all 100 projects the author has participated within the last 10 years of his professional experience, he has observed several different companies with several different organizational contexts (organization structure, business flows, employees and etc.). One of the most important aspect, he has recognized has been the resistance to change by some certain types of system users towards a new system. One unusual fact was that this negative behavior was observed, in some contexts, even before the very first tests of the new potential system by real users. The same fact for these users sustained after the very first tests. In some cases the attitude became even more negative after first trials whereas in some cases the attitude became more positive. In some very strange situations these users, who prior to using the system had negative attitudes, refused the opportunity of interacting with the system as much as possible.

Concluding, the author planned examining the above mentioned users' attitude and their perception of usability on a scientific platform. Is attitude affecting the perceived level of usability of a system? Do users with negative attitude towards computers perceive lower levels of usability? The answers to these questions will supply great information and will be of great use. If attitude and usability is somehow related then the professionals within the industry, prior to marketing and then developing a system, can analyze the customers' potential system users in regard to their computer attitudes. Found information can be shared with top-level managers of the potential customers and if the potential users of interest are found to have negative computer attitudes then the IT Company can be cautious. The IT Company can warn top-level managers that in the existence of users with negative computer attitudes, the perceived level of usability of a system can be lower. The company therefore can choose to come up with objective performance measures which will be selected as the success criteria for evaluation of the final system. An alternative can be to develop a program which will improve attitudes of these users who have negative attitudes towards computers so that when the real system is deployed perceived usability will be higher.

Chapter 2

Literature Review

2.1 Literature Review Regarding Usability

In this part of the thesis, an in-depth examination of different definitions and approaches to the concept of usability is supplied. A brief comparison of these various definitions is also existent. In the following sub sections of the literature review of usability, different techniques developed for measuring the concept is also analyzed with an in-depth examination of questionnaires that have been developed during the last 20 years.

2.1.1 Introduction to Usability / What is Usability?

Usability is an important goal to achieve in all aspects of product design. It is one of the most important criteria of the process. We have heard the following statement, in a variety of products by several different users, numerous times in our lives; "It is really difficult to use this product…" There are products in the market today even with their market launch they are a hit in a single day and there are some products does not matter how much their manufacturers invest upon promoting it is still not welcomed by potential customers. An example of the prior is Apple's iPod which became a market leader in less than a year all around the globe. Its success was in its design in the form of both elegance and even to a more important level, the ease to use the product by customers. Being able to operate almost any functionality with its innovative rotating wheel, the product is easy to use under almost every scenario (jogging, running, at bus, at car….). It is quite easy to observe that even a first time user is able to operate a significant percentage of the available functionality of the product in a couple of minutes. The physical usability characteristics of the product is subjectively perceived as easily usable by the customer and results in a highly welcoming, positive attitude towards the product. An example of the latter is Linux operating system (an operating system for PCs) which is in some cases freely available or available at a price which is significantly lower than its competitor, Microsoft Windows. Due to the reason that almost all pc users are acquainted with Microsoft Windows operating systems for a much longer period and perceived it as the reference when it comes to the ease of use, the Linux, with its less user friendly operability is perceived in the market as a product which “Technical People” use but not suitable for regular users. Does not matter whether truly the Windows is more user-friendly, still the perception of “User friendliness” of the product translates directly into a share which almost rules the entire market and is a norm. Concluding usability is a criterion which directly impacts a product’s / service’s success.

Can we come up with a universally accepted set of attributes / dimensions which constitute the concept of usability? One can quite easily answer this question by examining a couple of different scenarios, which are context dependent upon users. As a first one let’s take into account a hardware technical service provider in a software system. For the technical operators of the service provider a more usable, and preferred system is one which is easy to maintain, does not produce errors quite frequently, and it is easy to backup and restore the whole system from a previous functioning backup. This view of the technical operators does take into account their expectancies. A highly usable system for them is one which does not make them work for more hours to satisfy their daily routine jobs. Let’s switch to another users’ view that are interacting with the very same system as bank operators who are maintaining customer accounts on this very system. For them this system is usable if it is easy to find customer accounts within the database, if it is easy to examine past interactions with the bank, if it is easy to create new customers with not that much of effort and if help in the case of a failure is available with ease and etc. As can be seen when switched to a different user, the perception of “Usability” completely differs. Adding to the complexity, let’s take into account a complete different user and its view. Mark Dickinson is this bank’s top-level manager. He has to come up with a project which will satisfy the board of directors and in return announced to the shareholders as an important threshold in improving the performance of the bank and thus will boost the shares in the market. For him a highly usable system is one which is the cheapest that satisfies the expected set of functionalities (initial investment is low, and projected return is high, a better IOR), which is the cheapest in maintenance (variable costs) and it should be possible the custom develop new functionalities on to the software package with the internal IT department (internal code quality, documentation, modularity and other technical product properties are of importance, etc,.). Mark and his view of a highly usable system brought cost into the picture as a criterion of consideration. This was not a criterion of decision for the bank operators nor was it for the technical operators. The software interface and its abilities were not important for the technical operators nor was it for the top-level manager, Mark Dickinson but maybe the most important criterion for bank operators who are maintaining customer accounts. Concluding, the concept of usability and its dimensions differ up to a very significant degree depending upon different scenarios which is to be analyzed in-depth in the following pages.

When we focus in to the computers and software, usability is a topic which is highly popular and thus is of great discussion. Due to the very different scenarios of interacting with computer software and systems, there are different users, different views, different hardware setups and other variables which almost make every specific scenario entirely unique. Therefore comparison of findings of different authors under different field / laboratory conditions is impossible. Even the literature still lacks a universally agreed definition of the concept usability.

2.1.2 Different Definitions of the Concept Usability

The term usability was originally derived from the term “user friendly”. But due to the reason that this term had acquired significant subjective connotations “usability” was suggested to replace it (Folmer and Bosch 2004). Some authors defined this new term as a software quality (a narrow view) where as some defined it as an overall system design objective where usability is a goal to achieve (as in the case of ISO

9241 (ISO 9241-11, 1994), a broader view). Thus the authors with the latter view preferred to name the term usability as “quality in use” (Bevan suggested this term Bevan, 1995b).

When considered the existing literature, it is not easy to find a single universally acceptable definition of this concept. One can quite easily figure out how complicated this can get by understanding that although several authors has worked upon usability for more almost two decades, even now there is lack of consensus in this field. Still there are several authors and international bodies whose definition of usability is widely recognized such as; Shackel (Schakel, 1991), Nielsen (Nielsen, 1993a), ISO (ISO 9126 (ISO 9126, 1991) and ISO 9241 standards). Other definitions of usability which are not to be considered in-depth within this study can be found within the following books & studies: Constantine and Lockwood (Constantine and Lockwood, 1999); Hix and Hartson (Hix, Hartson, 1993); Preece (Preece, Rogers, Sharp, Benyon, Holland, Carey, 1994); Shneiderman (Shneirderman, 1986), Wixon and Wilson (Wixon, Wilson, 1997). The following statement successfully explains the current condition regarding the lack of consensus amongst different authors on defining usability:

> Although there is a consensus about the term usability, there are many different approaches to how usability should be measured; hence usability is defined in such a way as to allow these measurements. This definition has resulted in different definitions of usability, because authors have different opinions on how to measure usability. (Folmer and Bosch 2004)

Actually Folmer and Bosch’s statement conflicts within itself. The authors state that there is consensuses on the term usability but later add that every author in the field has different opinions regarding to what constitutes usability, the dimensions of usability. Thus if every author in the field has different opinions regarding the dimensions of usability then they do in fact have different definitions of usability meaning they do no have a consensus on the term. Thus the overall statement of “There is consensus about the term usability” is falsified. As a result when examined the different definitions of usability, the dimensions of the concept for the specific definition will also be provided to supply in-depth view of the author next to his / her definition of the term.

2.1.2.1 Schakel's Definition of Usability

Schakel is amongst the first to study usability and usability engineering. His definition of usability is as follows:

> The usability of a system is the capability in human functional terms to be used easily and effectively by the specified range of users, given specified training and user support, to fulfill the specified range of tasks, within the specified range of scenarios. (Schakel 1991)

This definition takes into account two sides of usability; the relative side and the objective side. The relativity of his definition states that usability is affected by the scenario and users. The objective side states that a system is usable if it is capable of effectively performing to fulfill the specified range of tasks by the users. So if it is possible to fulfill the tasks then the system has high usability. Although Schakel's definition clearly states the relativistic behavior of the concept, there is no guidance on how to measure it. This relativistic property of the concept is termed as its context-dependence meaning that any usability measurement is depending upon the user, environment, system and etc. Therefore user's properties (such as culture, attitude towards computers) or environmental conditions (lightning, noise... etc.) or system conditions (such as a low hardware performance, for the evaluation of the usability of a software interface) all impact the perceived level of usability. The context-dependent nature of the concept therefore adds to the complexity with subjective evaluations and perceptions of the user. Figure 2.1 displays his complete model where usability is a construct under product acceptance.

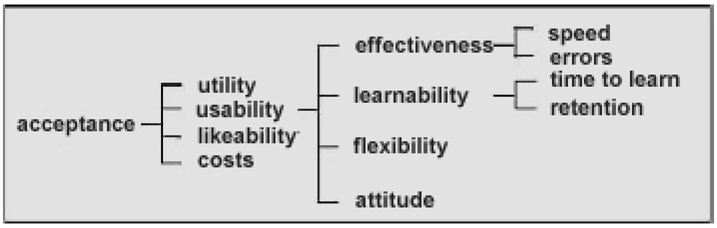


Figure 2.1: Schakel's Product Acceptance Definition & Dimensions

In Schakel's model product acceptance is the highest concept utility, usability, likeability and costs, are its constructs. Usability Construct is of composed of effectiveness, learnability, flexibility and attitude sub-constructs. When compared to other important usability definitions following overlaps are found: Effectiveness is also a criterion in ISO 9241-11, learnability is a criterion in ISO 9126. It is interesting to see that although Schakel mentioned effectiveness under usability, he has not taken efficiency as a construct to usability.

2.1.2.2 Nielsen's Definition of Usability

Nielsen is also amongst to first to study the concept of Usability. He does not give a precise definition of usability as Schakel does but rather examines it under the concept of product acceptance / acceptability (Figure 2.2) (Nielsen 1993a). His work regarding is as follows:

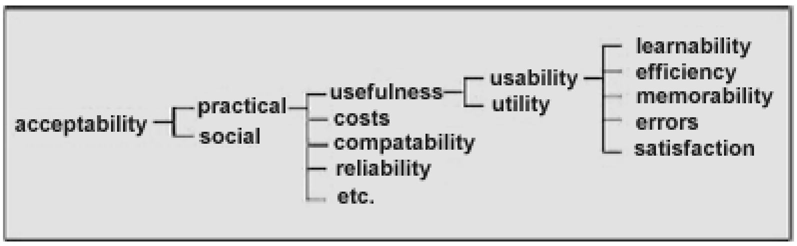


Figure 2.2: Nielsen's Product Acceptability Definition & Dimensions

He has placed usability under the usefulness dimension which is a sub dimension of practical acceptability of a product. The explanation of the concept takes into account 5 different dimensions: learnability, efficiency, memorability, errors and satisfaction.

- Learnability: Systems should be easy to learn. Users can rapidly start getting come work done with the system.
- Efficiency: Systems should be efficient to use. When a user has fully learned the system, productivity, will be possible on a high level.
- Memorability: Systems should be easy to remember, making it possible for casual users to return to the system after some period of not using the system, without having to learn everything all over again.
- Errors: The system should have a low error rate, which enables users to make few errors during the use of the system. When they do make errors they can easily recover from them. Catastrophic errors should not occur.
- Satisfaction: The system should be pleasant to use; which makes users subjectively satisfied when using it.

This is a real clear definition of the concept. All dimensions offered by Nielsen are system outputs meaning all are performance criterion measured when the real world system is deployed under real world scenarios. This is consistent in terms of being able to measure all dimensions that are offered in the same time horizon (cross-shot measurement) of system interaction.

As Schakel, Nielsen's approach also embodies both objective and subjective criterias of evaluation. Learnability, Efficiency, Memorability and Errors are the objective criterias of evaluation while satisfaction is a subjective one.

When compared to other important usability definitions following overlaps are found: Learnability is also a criterion in Schakel's and ISO 9126's definition.

Satisfaction only overlaps with ISO 9241-11's definition. Efficiency overlaps with only ISO 9241-11. Nielsen's and Schakel's Error dimension also overlaps (which is an effectiveness criterion).

In direct comparison to Schakel, Nielsen's explanation of the concept is more detailed and more closely resembles ISO 9126's approach.

2.1.2.3 ISO 9126's Definition of Usability

ISO has published several standards that focus upon usability, amongst them, ISO 9126, ISO 9126-1 and ISO9241-11 are most popular by practioners. ISO 9126 (1991) gave a definition of usability as follows; "Usability is a set of attributes of software which bear on the effort needed for use and on the individual assessment of such use by a stated or implied set of users." (ISO 9126, 1991)

This definition takes into account the subjective perspective by the assessment of the "implied set of users". The term "implied set of users", to a certain degree, reflect the very context dependent nature of the concept which is in harmony with the works of Schakel. Up to here, the ISO 9126's definition is consistent with the current views of the authors that worked upon this field. However, "Usability is a set of attributes of software" clearly limits the practioners by downgrading the concept into the properties of the software only. The overall performance perceived is not only limited to the software which is an essential but not the single sub-component of a system (IT-System in this context). Also adding to this very fact, the context dependency which is evident in this definition is only limited to the user however, other factors such as environment, training, and scenario are not taken into account. As a result of these limitations the view of original ISO 9126 has changed considerably throughout the time and transformed it self into an overall quality model by ISO 9126-1 (2000). In this quality model usability plays two different roles; an overall design objective and a component of software quality.

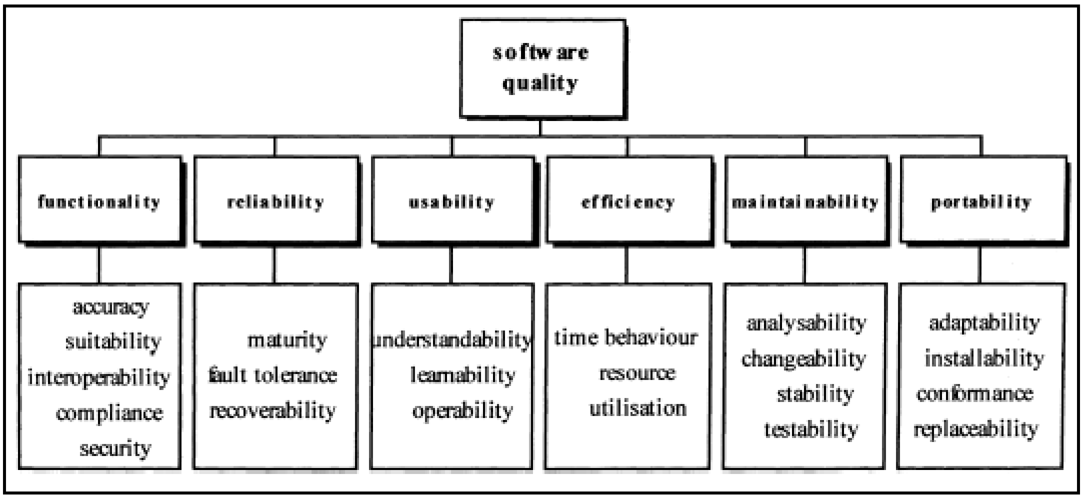


Figure 2.3: ISO 9126-1 Software Quality Model

In ISO 9126-1 software quality is referred to as "quality in use" and quality in use is defined as; "The capability of the software product to enable specified users to achieve specified goals with effectiveness, productivity, safety and satisfaction in a specified context of use." (Figure 2.3) (ISO 9126-1, 2000)

The definition of the ISO 9126-1 reflects the very welcomed perception of the concept today. This definition is almost identical to ISO 9241-11 definition except that it adds a further dimension into consideration which is "safety". According to Bevan the two definitions of ISO (referring to ISO 9126-1 and ISO 9241-11) are complimentary (Bevan, 2001).

The view of ISO 9126-1 encompasses two different roles of usability. These roles are:

- Product oriented role: Usability is a part of software design process. This is in harmony with the view of the original ISO 9126 (1991). The concept is an attribute of software quality.
- Process oriented role: Usability is a final goal; it is a design objective of the overall software system; the software should meet user needs.

The latter role; “Process oriented role” is referred to as “quality in use”. This is the broad definition of the concept in contrast with the narrow view of the concept where usability is a software attribute in product oriented role. ISO 9126-1 combined the definitions of ISO 9126 (product oriented role) with ISO 9241-11 (process oriented role).

Effectiveness, productivity and safety are objective dimensions offered by this definition where as satisfaction is the subjective one. The context dependence of the concept is also evident by the terms “Specified users”, “Specified context of use” and “Specified goals”. On the dimension of context dependence we find out that ISO 9126 only takes into account “users” where as the successor ISO 9126-1 takes into account also “context of use” and “specified goals”. Thus the ISO 9126-1’s context dependence is more detailed and satisfactory. One can quite easily come up with the conclusion that every usability study is unique due to the reason that the objectives of the system, the environment (physical, the organization) and the users are almost different in each and every scenario. The following are the dimensions of quality in use according to ISO 9126-1 model;

- Functionality: The capability of the software to provide functions which meet stated and implied needs when the software is used under specified conditions.
- Reliability: The capability of the software to maintain its level of performance when used under specified conditions.
- Usability: The capability of the software to be understood, learned, used and liked by the users, when used under specified conditions.
- Efficiency: The capability of the software to provide the required performance, relative to the amount of resources used, under stated conditions.
- Maintainability: The capability of the software to be modified. Modifications may include corrections, improvements or adaptation of the software to changes in environment, and in requirements and functional specifications.

- Portability: The capability of software to be transferred from one environment to another.

The following are the dimensions of usability which are offered by ISO 9126-1 (2000);

- Understandability: The capability of the software product to enable the user to understand whether the software is suitable, and how it can be used for particular tasks and conditions of use.
- Learnability: The capability of the software product to enable the user to learn its application.
- Operability: The capability of the software product to enable the user to operate and control it.
- Attractiveness: The capability of the software product to be attractive to the user. For instance the use of colors or nature of graphical design.

What is most impressive about the redefined and enriched view of the ISO 9126 standard is that (by ISO 9126-1, 2000) it meets the very different definitions required due to their different processes and ways interactions of two different sets of users in a single definition and international standard. These sets of users are; software engineers and the final / end users of the system. The software engineers mostly interact with the system in the design phase where the product is far from market launch and is even on the “drawing board”. At this time of the life-cycle of a software, product engineers are trying to design how the final users will interact (interface design) with the product and will judge whether it is understandable, learnable, operable and attractive to them to a satisfactory degree in contrast with the resources utilized to realize the whole system. For engineers a highly usable system, from their perspective and their objectives is mostly one which is easily portable to different running environments (mobile gadgets, different running platforms such as Unix based, Windows based servers in a scenario of client and server architectural software), can be easily modified to add further functionalities and corrections are easy to apply. However, these criteria are almost non-decisive when we switch to the views of the final / end user who pay more attention to the overall perceived quality of the system. For them the most obvious criterion are functionality, reliability, usability and to efficiency. ISO 9126-1 is capable of defining usability satisfactorily for both of these types of users.

2.1.2.4 ISO 9241's Definition of Usability

ISO 9241-11 (1994) is amongst several standards that are developed for the purpose of standardizing the way of interaction of a user with computers (Figure 2.4). Chapter 11 of ISO 9241 (Guidance on Usability), focuses on the ergonomic requirements for office work with Visual Display Terminals. This standard has led to guidelines for software interfaces based on research by Macleod (1994) and Bevan (1995a). According to ISO 9241-11 (1994) usability is defined as; "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."

This definition takes into account objective criterion such as effectiveness and efficiency. Satisfaction is the only subjective criteria suggested. When remembered ISO 9126-1's definition of the concept, one can quite easily find out that both definitions are identical except that ISO 9126-1 adds safety. Thus ISO 9241-11 is the prior standard in which usability is defined up to a real satisfactory level which is widely accepted and welcomed by practioners. It is in harmony with the broad view of the concept which clearly states that usability is a design objective of the overall system which is a final goal itself. This broad view is identical to the product-oriented role of the ISO 9126-1.

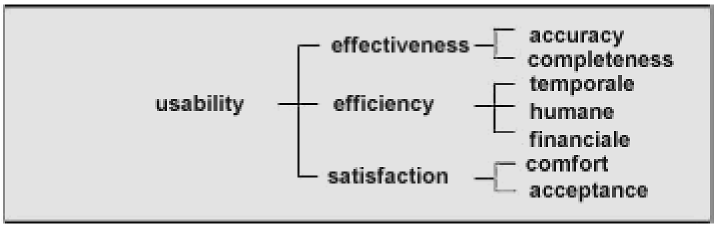


Figure 2.4: ISO 9241-11 Usability Model

- Effectiveness: Measures of effectiveness relate the goals or sub goals of the user to the accuracy and completeness with which these goals can be achieved.
- Efficiency: Measures of efficiency relate the level of effectiveness achieved to the expenditure of resources. Relevant resources can include mental or physical effort, time, materials or financial cost. For example, human efficiency could be measured as effectiveness divided by human effort, temporal efficient as effectiveness divided by time or economic effiency as effectiveness divided by cost.
- Satisfaction: Satisfaction measures the extent to which users are free from discomfort, and their attitudes towards the use of the product.

ISO 9241-11 provides the following information regarding the dimensions of the concept. Effectiveness is composed of accuracy and completeness. The user should be able to complete his / her goals with accuracy. Efficiency is composed of temporal, humane and financial. This wide view of efficiency takes into account different views of different actors within the organization who are responsible of making the strategic decision to utilize this software product, such as the top-level manager who is mostly concerned with financial efficiency or the actual system user (operator) who is mostly concerned with time efficiency. Satisfaction is composed of comfort and acceptance. One can quite easily refer back to the definitions of Nielsen and Schakel who takes usability as a dimension to the concept of product acceptance and see the correlation with ISO 9241-11's acceptance dimension.

The framework of usability according to the ISO 9241-11, which pays much importance to the context dependence (context is composed of user, task, equipment and environment) of the concept, is as follows:

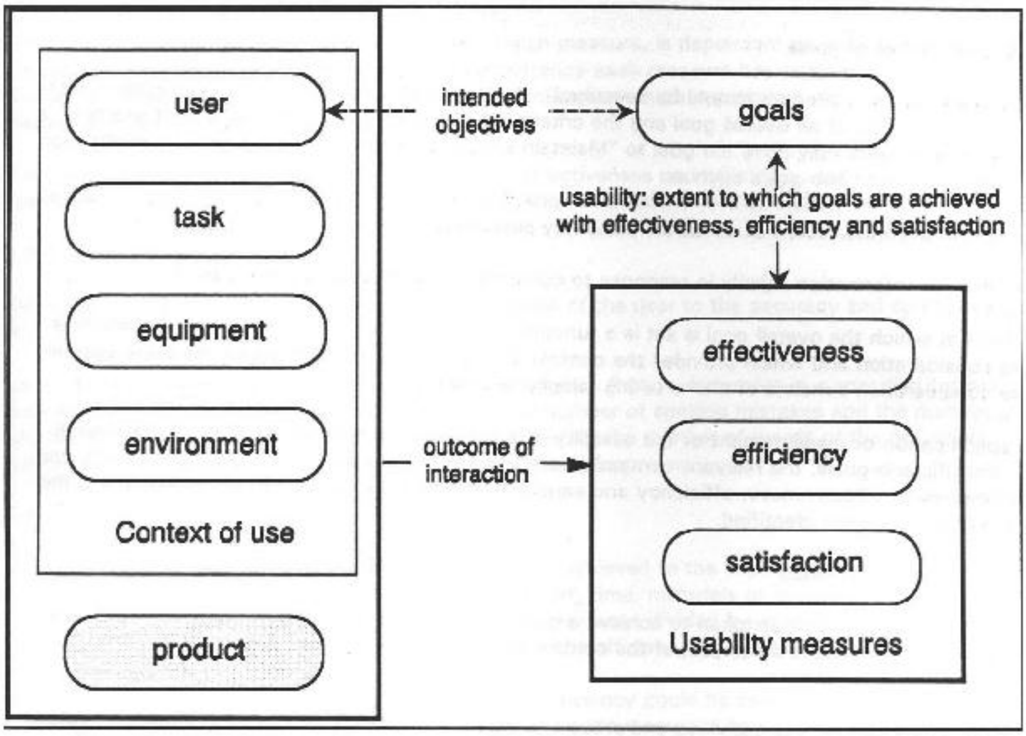


Figure 2.5: Joined Display of ISO 9241-11 and ISO 9126-1 Models

The definition of the concept pays much importance to the context dependence. The context dependence takes into account specified users, specified goals and specified context of use. Therefore any usability definition or measurement of a system in evaluation is dependent to a very important level to the actual users, the goals that are to be achieved, the practical and social environment of usage and other context factors. The following paragraph from the standard clearly explains the level of importance paid to this aspect:

> ISO 9241-11 emphasizes that visual display terminal usability is dependent on the context of use and that the level of usability achieved will depend on the specific circumstances in which a product is used. The context of use consists of the users, tasks, equipment (hardware, software and materials),

> and the physical and social environments which may all influence the usability of a product in a work system. Measures of user performance and satisfaction assess the overall work system, and, when a product is the focus on concern, these measures provide information about the usability of that product in the particular context of use provided by the rest of the work system. The effects of changes in other components of the work system, such as the amount of user training, or the improvement of the lightning, can also be measured by user performance and satisfaction. (ISO 9241-11, 1994)

As evident by the above explanation, the usability measurements are context oriented by their nature. It is not possible and wise to compare different measurements of the same particular software product in different organizations as the overall user performance and satisfaction is effected to a very certain degree by the different contexts of utilization. Therefore usability measurements are most useful in scenarios where the objective is to measure the relative utility gained by a single or a tiny improvement utilized within the system in almost the exact identical context. The standard clearly emphasizes against the comparison of different usability measurements directly against each other.

> Care should be taken in generalizing the results of any measurement of usability to another context which may have significantly different types of users, tasks, or environments. If measures of usability obtained over short periods of time, the values may not take account of infrequent events which could have a significant impact on usability, for example intermittent system errors. For a general purpose product, it will generally be necessary to specify or measure usability in several different representative contexts, which will be a subset of the possible contexts of the tasks which can be performed. There may be differences between usability in these contexts. (ISO 9241-11, 1994)

The standard supplies in-depth information for a usability professional to identify relevant information in a specific context of use for measuring effectiveness, efficiency and satisfaction.

2.1.3 How Different Definitions Overlap Each Other?

Figure 2.6 taken from Folmer and Bosch (2004) is a good indicator of the overlap between different definitions of usability. However one should be conscious as to the possibility of the variations of different definitions for the proposed dimensions of the concept (I.E learnability of Schakel focuses on "time-to-learn" and "retention" aspects however Nielsen's definition of the very same learnability dimension is not that clear and thus there could be differences). It should also be noted that on the following figure it is not entirely correct to directly compare ISO 9241-11's view of usability (broad view of usability) with the narrow views of usability of the other authors and / or standards (explained in details later).

E-Folmer, J. Bosch / The Journal of Systems and Software 70 (2004) 61-78
Overview of Usability Definitions

	Schakel	Nielsen (1993)	ISO 9241-11	ISO 9126
User performance (objective)	Learnability-teim to learn	Learnability		Learnability
	Learnability-retention	Memorability		
	Effectiveness-errors	Errors	Effectiveness	
	Effectiveness-task time	Efficiency	Efficiency	
				Operability
				Understandability
	Flexibility			
User view (subjective)	Attitude	Satisfaction	Satisfaction	Attractiveness

Figure 2.6: Overview of Usability Definitions

On a general bases all authors examine usability on two dimensions:

- Objective Criterion: A fact which is measurable by some objective data such as time to complete the task, accuracy of the action, time to learn the software and etc,.
- Subjective Criterion: A fact which is subjective in nature such as satisfaction with the usage of a software or finding the software attractive (i.e. the interface of the software; graphics, icons, texts…)

Learnability (exists in 3 of 4 definitions), effectiveness (exists in 3 of 4 definitions), efficiency (exists in 3 of 4 definitions) and satisfaction (exists in all definitions if finding the software attractive and having a good positive attitude software is to be taken into account as satisfaction) are the most popular dimensions.
ISO 9241-11 paved the way to the redefined ISO 9126-1 (2000) standard to evaluate usability from a broader view (process oriented role). ISO 9126-1 (2000) is the only definition which joins two different views on usability in a single framework of evaluation. These views are: the software engineers view (product oriented view) and the end / final users' view (quality in use: usability is a final goal to achieve process oriented role).

ISO's definitions are complimentary to each other. However when comparing these two standards one should clearly state which different view of ISO 9126-1 is being taken into account due to the reason that ISO 9126-1's product oriented view of usability is not the same thing as ISO 9241-11's process oriented broader view. This important consideration should also be paid important attention when ISO 9241-11's usability definition is to be compared and / or analyzed in contrast with Nielsen's and Schakel's definitions. Actually ISO 9241-11's view of usability (broad view of usability) can be said to be the same construct with product acceptance of Schakel and product acceptability of Nielsen. In this view of the concept; the usability of the overall system / software is the final goal; it should be effective and efficient on an objective basis and users should be satisfied with the overall usage thus meaning that they should be accepting the product as a good solution to their problem domain (as in terms of Nielsen's and Schakel's definitions in which product acceptance / acceptability is the final goal). On another view when we focus onto usability on a narrower view such as the one in ISO 9126-1's product oriented approach (usability dimension of the overall quality in use) Nielsen's usability dimension which is a sub-dimension of usefulness and Schakel's usability dimension which is a sub-dimension of product acceptance must be taken into account. As a summary; one should not conflict between usability on a narrow view (which is actually the usability of the software and in many instances this means interface usability) and usability on a broader view (usability of the overall solution with its software, hardware, costs, contexts…) and when comparing different definitions a statement regarding to which view is considered should be made clearly. This is the exact reason why we commented upon the figure taken from Folmer and Bosch (2004). In that figure the narrow view of Nielsen, Schakel and ISO 9126-1 is on comparison with broad view of ISO 9241-11. There is no the so-called "narrow view" of usability in ISO 9241-11.

Delving deeper into different definitions and views of the authors in the field, one sees important problems. The ambiguity of the term resulted different views of usability by different authorities. There is overlapping and interference but lack of consensus. Adding to the complexity level, all authors proposed several dimensions which they claim to explain the concept but even on a dimension basis, the explanations of these dimensions differ thus one can not be sure that learnability of Nielsen means / measures the very same aspect as of learnability of ISO 9126-1. Finally the context dependence nature of the concept makes it impossible to compare different usability tests directly and thus each and every observer in the field tries to improve their perception of usability of a system on their own isolated level (testing, measurements and etc,.). Summing up, all these limitations force us to use usability evaluation / testing methods / tools to investigate relative improvements on a system level and their impact upon relativistic usability of specific context driven scenarios. This is the exact reason why latter definitions of usability, especially the ones by ISO, pay significant attention to context dependence. Both ISO 9241-11 (1998) and redefined ISO 9126-1 (2000) suggests that the context is to be clearly stated with any measurements of usability.

2.1.4 Measuring Usability

The literature regarding usability is full of several different kinds of methods developed for measuring the concept. Some of these methods try to predict / forecast a system's usability within the design phase and some try to evaluate an already existing live system for future improvements. Regarding to the stage within the life cycle of the system the methods which could be utilized greatly differ. The categorizations of these different methods are also a discussion. Taking into account

Zhang's (2001) view usability testing / measurement methods can be classified into three distinct categories. These are:

- Usability Testing
- Usability Inspection
- Usability Inquiry

2.1.4.1 Usability Testing

This approach requires representative users to complete and test-drive the system on typical tasks that will be performed by the system. The users can work on semi-finished systems / products, therefore it can be deployed within specific stages of the system design cycle for recognizing early usability flows and deploying improvement plans as soon as possible before it gets to costly to redesign the solution. The evaluators evaluate the results of the users for deciding upon improvements. This testing method is especially suitable for measuring the interface usability of a system (There for it measures the narrow view of usability (interface usability) not the broad view, as existent in ISO 9126-1, of usability). Some methods which are utilized within this classification are;

- Coaching method (Nielsen, 1993a)
- Performance measurement (Nielsen, 1993a; Soken, Reinhart, Vora, Metz, 1993)
- Remote Testing (Hartson, Castillo, Kelso, Kamler, Neale, 1996)
- Teaching method (Vora and Helander, 1995)
- Thinking aloud protocol (Nielsen, 1993a)

2.1.4.2 Usability Inspection

This approach requires the availability of specialists, experts. These experts examine the system and judge whether the user interface follow established usability principles (such as colors within the interface, help dialog boxes, error messages, interaction via a mouse and keyboard, etc.). This approach does not require the availability of representative and / or actual system users. As usability testing, this classification is also focused on the narrow view of usability and the methods available are convenient for utilization while the system is within the design stage. Some methods which are commonly utilized under this classification are:

- Heuristic evaluation (Nielsen 1994b).
- Cognitive walkthrough (Wharton, Rieman, Lewis, Polson 1994).
- Perspective-based inspection (Zhang, Basili, Shneiderman 1998).
- Standards inspection / guideline checklists (Wixon , Jones, Tse, Casaday 1994).

2.1.4.3 Usability Inquiry

This approach requires the availability of the actual finished system and the actual users interacting with the system. The evaluators try to gather information / actual feedbacks from the actual users from their actual perception of the relative usability of the system. The gathered data can be both objective and subjective. An example to an objective criterion is; “Time to complete the task” which is quantifiably measurable and an example to a subjective criterion is; “The level of satisfaction”. This classification can both measure the narrow view and / or broad view of usability regarding to the testing approach and the data gathered. Some of the methods that are deployed in this approach are:

- Field observation (Nielsen 1993a).
- Interviews / Focus Groups (Nielsen, 1993a).
- Surveys (Alreck and Settle 1994).
- Logging Actual Use (Nielsen 1993a).
- Proactive field Study (Nielsen 1993a).

The inquiry methods are widely utilized within the industry. The methods mainly aim to measure the experiences of actual users in a real world scenario and use the gathered data for improving the system for a better experience. Therefore they are mainly used for measuring the improvements and their impact upon perceived level of usability of a specific system. Different measurements of different systems can not be compared against each other as each system is a complete different scenario that is composed of different goals, objectives, hardware, users, environment and etc.

One should clearly understand that for a software solution these three classification of usability measurement are used accordingly in each different stages of the software life-cycle (i.e.; usability testing and / or inspection is utilized while the system is being developed (design-stage), usability inquiry methods are used when the actual system is developed and potential improvements are analyzed and / or investigated). Within inquiry methods questionnaires are the most popular. There are significant numbers of questionnaires developed within the literature. All these questionnaires are psychometrically evaluated by different authors throughout the time. Most popular ones are:

- QUIS: Questionnaire for user interface satisfaction (Chin, Diehl, Norman, 1988).
- PEUE: Perceived usefulness and ease of use (Davis 1989).
- NHE: Nielsen’s heuristic evaluation (Nielsen 1993a).
- NAU: Nielsen’s attributes of usability (Nielsen 1993a).
- PSSUQ: Post-study system usability questionnaire (Lewis 1992a).

- CSUQ: Computer system usability questionnaire (Lewis 1995).
- ASQ: After scenario questionnaire (Lewis 1995).
- SUMI: Software usability measurement inventory (HFRG 2002).
- MUMMS: Measurement of usability of multimedia software (HFRG 2002).
- WAMMI: Website analysis and measurement inventory (HFRG 2002).
- EUCSI: End user satisfaction instrument (Doll, Torkzadeh 1994).

All these questionnaires are developed upon the authors' view and favored usability definition and thus one should be cautious and pay attention to the authors' view of usability before using them. This is the exact reason to why each usability measurement pays significant amount of time to first explain their favored definition of usability first and then detail the measurement method, instrument and etc.

Zhang's (2001) classification of usability is not the only classification available. Another good classification of usability measurement methods is by Gediga, Hamborg and Düntsch (?). They classify usability evaluation techniques into two distinct categories:

- Descriptive Evaluation Techniques: They are used to describe the status and the actual problems of software in an objective, reliable and valid way. These techniques are user based and can be subdivided into several approaches:
    - Behavior based evaluation techniques record user behavior while working with a system which "produces" some kind of data. These procedures include observational techniques and "thinking-aloud" protocols.
    - Opinion based evaluation methods aim to elicit the user's (subjective) opinions. Examples are interviews, surveys and questionnaires.
    - Usability Testing stems from classical experimental design studies. Nowadays, Usability testing (as a technical term) is understood to be a combination of behavior and opinion based measures with some

amount of experimental control, usually chosen by an expert. Observe that all descriptive evaluation techniques require some kind of prototype and at least one user. Note furthermore that the data gathered by a descriptive technique need some further interpretation by one or more experts in order to result in recommendations for future software development.

- The predictive evaluation techniques have as their main aim to make recommendations for future software development and the prevention of usability errors. These techniques are expert or at least expertise – based, such as Walkthrough or inspection techniques. Even though the expert is the driving power in these methods, users may also participate in some instances.

  Note that predictive evaluation techniques must rely on "data". In many predictive evaluation techniques, such "data" are produced by experts who simulate "real" users. The criteria objectivity and reliability, which are at the basis of descriptive techniques, are hard to apply in this setting. Because validity must be the major aim of evaluation procedures, there are attempts to prove the validity of predictive evaluation techniques directly, e.g. by comparing "hits" and "false alarm" rates of the problems detected by a predictive technique

It is easy for one to see that Gediga, Hamborg and Düntsch's "Descriptive evaluation techniques" classification covers "Usability Inquiry" of Zhang's (2001) classification. As such their predictive evaluation techniques cover "Usability Testing" and "Usability Inspection" of Zhang's (2001) classification.

Depending upon the objective of any research study either a single or a combination of the above mentioned classifications can be deployed. If one's approach is to develop guidelines then usability testing and / or inspection methods which mainly require the availability of experts are more appropriate. If one's approach is to examine a specific context (system, user, goals, environment and etc,.) and detect usability problems for future improvement plans then usability inquiry methods are more appropriate. Depending upon the specific techniques of interest objective and / or subjective data is to be gathered. It should also be pointed out that each technique is to be utilized and fine-tuned to the specific scenario of testing. Another important aspect is that, one should clearly state the definition of interest before making any attempts to measure usability of a specific system.

When one's aim is to examine the impact of usability upon an already existing system for examining correlations with specific variables of interest and / or for future improvement recommendations, the inquiry methods comes into play. Especially questionnaires in inquiry methods are of great interest within the literature. These measurement tools are easy to apply (in contrast to the methods which require specific hardware setups for data gathering such as the one in: "Physiological responses to different web page designs" (Ward, Marsden 2003), this article examines one's heart rate and other physiological parameters and their variability against different web page designs) and are good at capturing subjective data from users. It is important to assess subjective data such as satisfaction on usability inquiry methods as at this level (the software is fully developed and actual users are interacting with the final system) as user satisfaction and their perceived level of the actual system is the real criteria regarding to the perceived performance of the final system. Therefore evaluating a system's performance solely on objective performance criteria such as time to complete a task is inappropriate as the users can resist using a system which they do not favor and / or dislike and / or dissatisfied although objective performance is on a high level.

Focusing onto questionnaires in usability inquiry classification of Zhang's (2001); one finds the following questionnaires mostly used in usability studies; QUIS (Chin 1988), PSSUQ (Lewis 1992a), CSUQ (Lewis 1995), ASQ (Lewis 1995), SUMI (HFRG 2002), MUMMS (HFRG 2002), WAMMI (HFRG 2002). We will focus onto each questionnaire and supply the details in the following headlines.

2.1.4.3.1 QUIS: Questionnaire for user interface satisfaction (Chin, 1988)

Chin's, 1998 work regarding user interface satisfaction is on comparing the relative perceived usability of a liked versus disliked software and a Command Line System (CLS, like old dos based environment) versus a Menu Driven Application (MAD, much like today's windows environment). The original work of Chin was conducted on 150 users. The questionnaire (QUIS) contains several semantic differential scale items. Examples to some of the bipolar adjectives are:

- Terrible, Wonderful
- Difficult, Easy
- Frustrating, Satisfying
- Dull, Stimulating
- Rigid, Flexible
- Confusing, Very Clear
- Inconsistent, Consistent
- Unhelpful, Helpful

The reliability of Chin's original work was 0.94 (Cronbach's alpha) (QUIS version 5.0, 21 items). Earlier versions of QUIS was a long questionnaire (there were several versions of QUIS, QUIS 3.0 consisted of 103 item ratings with a Cronbach's alpha of 0.94, QUIS 4.0 consisted of 70 items with a Cronbach's alpha of 0.89). The interitem alpha values of QUIS 5.0 ranged from .933 to .939. Further versions decreased the number of items significantly. Originally Chin tried to examine the discriminatory power of the questionnaire by examining the differences between questionnaire responses on liked versus disliked and CLS versus MDA software. Factor analysis of QUIS found 4 factors and Chin has named these factors as: Learning, Terminology and Information flow, System Output and System Characteristics.

The mean differences between liked and disliked software can be examined on each specific items in Table 2.1.

Means of Ratings for Like vs. Dislike Groups				
	Like		Dislike	
	Mean	St. Dev.	Mean	St. Dev.
A. Overall Reactions to the System				
1. (terrible/wonderful)	7.21	1.21	4.44	2.41 ****
2. (frustrating/satisfying)	7.12	1.43	3.29	2.14 ****
3. (dull/stimulating)	6.68	1.35	3.75	2.14 ****
4. (difficult/easy)	5.59	2.05	4.33	2.61
5. (inadequate power/adequate power)	7.00	1.52	5.06	2.73 **
6. (rigid/flexible)	6.28	1.76	3.52	2.53 ***
B. Screen				
1. Characters on the computer screen (hard to read/easy to read)	6.94	1.80	7.19	1.80
2. Highlighting on the screen simplifies task (not at all/very much)	6.20	1.81	6.12	2.50
3. Organization of information on screen (confusing/very clear)	6.29	1.62	5.76	1.89
4. Sequence of screens (confusing/very clear)	6.45	1.46	5.69	1.48
C. Terminology and System Information				
1. Use of terms throughout system (inconsistent /consistent)	7.09	1.42	6.68	1.62
2. Computer terminology is related to the task you are doing (never/always)	6.39	1.90	5.79	1.93
3. Position of messages on screen (inconsistent/consistent)	7.24	1.44	6.25	2.08
4. Messages on screen which prompt user for input (confusing/clear)	6.03	1.98	5.00	2.52
5. Computer keeps you informed about what it is doing (never/always)	6.24	1.89	5.24	2.46
6. Error messages (unhelpful/helpful)	5.97	2.15	4.71	3.24
D. Learning				
1. Learning to operate the system (difficult/easy)	5.67	2.10	3.53	2.67 **
2. Exploring new features by trial and error (difficult/easy)	5.62	2.00	3.76	2.46 **
3. Remembering names and use of commands (difficult/easy)	5.52	2.40	4.56	2.68
4. Tasks can be performed in a straight-forward manner (never/always)	5.94	1.72	4.65	2.50 *
5. Help messages on the screen (unhelpful/helpful)	5.94	2.11	4.94	3.01
6. Supplemental reference materials (confusing/clear)	5.29	2.00	3.93	2.74
E. System Capabilities				
1. System speed (too slow/fast enough)	6.09	2.28	4.29	2.61 *
2. System reliability (unreliable/reliable)	7.45	1.50	6.35	1.66 *
3. System tends to be (noisy/quiet)	6.50	2.05	7.14	1.66
4. Correcting your mistakes (difficult/easy)	6.64	1.92	4.71	2.44 **
5. Experienced and inexperienced users' needs are taken into consideration (never/always)	5.63	1.81	3.88	2.09 **
note: * denotes p<.05 ** denotes p <.01 *** <.001 **** <.0001				

Table 2.1: Mean of Ratings Table from a QUIS Analyses

In overall reactions liked software rated significantly higher results than disliked software. In 3 items this result was significant at the level of $p<.001$. Easy / difficult was the only item that produced a result which was not significant ($p<.05$). For the component questions (the latter 21 items under the headings screen, terminology and system information, learning, system capabilities) some of the items were significant at the level $p<.05$, none of the items were significant at the level of $p<.001$.

The mean differences between CLS and MDA software can be examined on each specific item in the figure in the next page. In general, all the MDA mean ratings were higher than CLS. All of the overall reaction items were significant on t-test mostly at the level of $p<.0001$. The single exception was difficult / easy. One can remember that this was the case for difficult / easy also on the comparison of liked versus disliked software. In 21 component questions, 8 were significant at the level of $p<.001$.

Generally, Chin’s QUIS 5.0 discriminated MDA versus CLS better than liked versus disliked software. Chin’s attempt to measure the user interface satisfaction in 1998 was amongst the first in the field. Chin, Diehl and Norman established a reliable questionnaire which was able to measure existent satisfaction of a system user on 4 factors. Therefore it proved important value for examining an existing product on 4 dimensions and pointing out the relative points of consideration which have to be focused for improving the product. It is a good tool to either compare different products amongst or a latter version of the same product with an earlier version for measuring the impact of modifications on satisfaction.

Means of Ratings for Command Line System (CLS) vs. Menu Driven Application (MDA) ps				
	CLS		MDA	
	Mean	St. Dev.	Mean	St. Dev.
A. Overall Reactions to the System				
1. (terrible/wonderful)	5.33	1.47	7.36	1.11 ****
2. (frustrating/satisfying)	5.07	1.96	6.84	1.60 ***
3. (dull/stimulating)	4.65	2.22	5.83	1.53 *
4. (difficult/easy)	4.59	1.58	5.24	1.56
5. (inadequate power/adequate power)	4.96	2.19	7.75	1.42 ****
6. (rigid/flexible)	4.33	2.17	6.88	1.54 ****
B. Screen				
1. Characters on the computer screen (hard to read/easy to read)	6.08	2.78	7.62	1.20 *
2. Highlighting on the screen simplifies task (not at all/very much)	5.00	2.90	6.72	1.81 *
3. Organization of information on screen (confusing/very clear)	4.36	2.08	7.40	1.29 ****
4. Sequence of screens (confusing/very clear)	5.18	1.72	7.20	1.10 ***
C. Terminology and System Information				
1. Use of terms throughout system (inconsistent /consistent)	6.42	1.89	7.54	1.06 *
2. Computer terminology is related to the task you are doing (never/always)	5.46	2.08	6.63	1.35 *
3. Position of messages on screen (inconsistent/consistent)	6.00	2.64	8.04	0.95 ***
4. Messages on screen which prompt user for input (confusing/clear)	4.77	2.25	6.44	1.58 **
5. Computer keeps you informed about what it is doing (never/always)	4.19	1.79	6.71	1.33 ****
6. Error messages (unhelpful/helpful)	3.54	1.92	5.80	1.61 ****
D. Learning				
1. Learning to operate the system (difficult/easy)	3.56	1.78	5.08	2.12 **
2. Exploring new features by trial and error (difficult/easy)	4.35	2.24	5.56	1.98 *
3. Remembering names and use of commands (difficult/easy)	4.48	2.17	5.04	2.30
4. Tasks can be performed in a straight-forward manner (never/always)	4.74	1.75	6.16	1.31 **
5. Help messages on the screen (unhelpful/helpful)	3.74	2.16	6.16	1.80 ***
6. Supplemental reference materials (confusing/clear)	4.30	2.28	5.84	1.60 **
E. System Capabilities				
1. System speed (too slow/fast enough)	5.31	2.31	6.84	1.34 **
2. System reliability (unreliable/reliable)	7.19	1.77	7.48	1.36
3. System tends to be (noisy/quiet)	6.33 1.88	7.13	2.01	
4. Correcting your mistakes (difficult/easy)	5.24	2.01	7.04	1.31 ***
5. Experienced and inexperienced users' needs are taken into consideration (never/always)	3.80	2.12	6.00	1.53 ***
note: * denotes p<.05 ** denotes p <.01 *** <.001 **** <.0001				

Table 2.2: Mean of Ratings Table from a QUIS Analyses (2)

QUIS 5.0 questionnaire of Chin (1988) can be examined in Figure 2.7. It contains 6 overall reaction items and 21 main component items under 4 factors.

OVERALL REACTION TO THE SOFTWARE		0	1	2	3	4	5	6	7	8	9		NA
1. Comments:	terrible											wonderful	
2.	difficult											easy	
3.	frustrating											satisfying	
4.	inadequate power											adequate power	
5.	dull											stimulating	
6.	rigid											flexible	
SCREEN		0	1	2	3	4	5	6	7	8	9		NA
7. Reading characters on the screen	hard											easy	
8. Highlighting simplifies task	not at all											very much	
9. Organization of information	confusing											very clear	
10. Sequence of screens	confusing											very clear	
TERMINOLOGY AND SYSTEM INFORMATION		0	1	2	3	4	5	6	7	8	9		NA
11. Use of terms throughout system	inconsistent											consistent	
12. Terminology related to task	never											always	
13. Position of messages on screen	inconsistent											consistent	
14. Prompts for input	confusing											clear	
15. Computer informs about its progress	never											always	
16. Error messages	unhelpful											helpful	
LEARNING		0	1	2	3	4	5	6	7	8	9		NA
17. Learning to operate the system	difficult											easy	
18. Exploring new features by trial and error	difficult											easy	
19. Remembering names and use of commands	difficult											easy	
20. Performing tasks is straightforward	never											always	
21. Help messages on the screen	unhelpful											helpful	
22. Supplemental reference materials	confusing											clear	
SYSTEM CAPABILITIES		0	1	2	3	4	5	6	7	8	9		NA
23. SYSTEM speed	too slow											fast enough	
24. SYSTEM reliability	unreliable											reliable	
25. SYSTEM tends to be	noisy											quiet	
26. Correcting your mistakes	difficult											easy	
27. Designed for all levels of users	never											always	
		0	1	2	3	4	5	6	7	8	9		NA

Figure 2.7: A QUIS Questionnaire Example

2.1.4.3.2 PEUE: Perceived Usefulness and Ease of Use (Davis 1989)

Davis's Perceived Usefulness and Ease of Use questionnaire mainly develops upon his Technology Acceptance Model (TAM) (Davis, 1986). TAM (Figure 2.8) makes it possible to evaluate the systems before they are purchased (evaluations on prepurchase) on a trial period therefore is an important and respected tool in Management Information Systems.

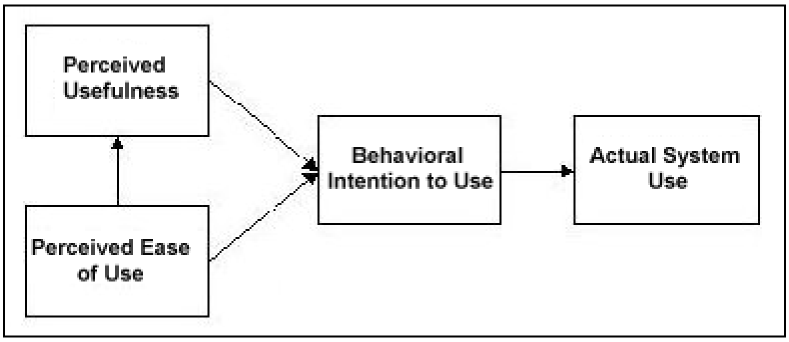


Figure 2.8: Technology Acceptance Model of Davis*

Its reliability and validity has been tested several times throughout the literature (William, Doll, Xiaodong 1998). Most of the research which focused on this matter focused on Davis's instruments ability to act as invariant between different contexts such as amongst users with different computer experiences, different organizations and different types of applications (such as word processing, spreadsheets, database applications, etc.). Davis's latter work showed that TAM was capable of explaining the %40 or variance in intentions to use of a specific application in a specific context (Venkatesh, Davis 2000). TAM focuses on two beliefs of system users, perceived usefulness and perceived ease of use. These two beliefs, according to TAM, will result and can predict computer acceptance behavior.

> TAM posits that two particular beliefs, perceived usefulness and perceived ease of use, can predict computer acceptance behavior. Perceived usefulness is defined as the prospective user's subjective probability that using a specific application system will increase his or her job performance within an organizational context. Perceived ease of use refers to the degree to which the prospective user expects the target system to be free of effort. Perceived usefulness and ease of use are meant to be fairly general determinants of user acceptance. Davis et al. (1989) described them as belief sets that are meant to be readily generalizable to different computer systems (applications) and user

* Taken from http://www.istheory.yorku.ca/Technologyacceptancemodel.htm 25.12.2007

> populations. They argued that the ability to take robust, well-formed measures of the determinants of user acceptance early in the development process (i.e., initial exposure data) can have an impact on system acceptance by enabling developers to weed out bad systems, refine the rest, and generally cut the risk of delivering finished systems that get rejected by users. (Hendrickson, Deng 1998)

Figure 2.9 displays an instrument that is deployed in a study which examines the social influences for TAM in a course delivery system (Shen, Laffey, Lin, Huang 2006). Perceived usefulness and perceived utility each is measured by six, 7-point Lickert scale items.

	Strongly Disagree			Neutral			Strongly Agree
	1	2	3	4	5	6	7
1. My **INSTRUCTORS** expect me to use <the course management system (CMS)>							
2. My **INSTRUCTORS** want me to use <the CMS> frequently							
3. Generally speaking I try to do what **INSTRUCTORS** think I should do							
4. The **ZONE MENTORS** expect me to use <the CMS>							
5. The **ZONE MENTORS** want me to use <the CMS> frequently							
6. Generally speaking I try to do what the **ZONE MENTORS** think I should do							
7. My **CLASSMATES** expect me to use <the CMS>							
8. My **CLASSMATES** want me to use <the CMS> frequently							
9. Generally speaking I try to do what **CLASSMATES** think I should do							
10. Using <the CMS> helps me learn about & accomplish the course requirements quickly							
11. Using <the CMS> helps me to be a productive student							
12. Using <the CMS> enhances my effectiveness on the course coursework							
13. Using <the CMS> makes it easy to do the course coursework							
14. Using <the CMS> improves my academic performance							
15. I find <the CMS> a useful tool for my learning in this course							
16. Learning to use <the CMS> is easy for me							
17. <The CMS> is flexible to interact with							
18. I find it's easy to get <the CMS> to do what I want to do							
19. It's easy for me to become skillful at using <the CMS>							
20. My interaction with <the CMS> is clear and understandable							
21. <The CMS> is easy to use							

Items Regarding PU, PE

Figure 2.9: A TAM Questionnaire

Davis's PEUE instrument is a great tool which is validated several times throughout the literature and has been found to be both reliable and valid. This instrument is convenient for being deployed in organizational contexts where one's aim is to choose amongst available Information System products and want to discriminate between ones which will be favored by potential users to their high perceived usefulness and utility. Therefore it is a tool which is to be used in decision making processes (i.e. in the case of corporate buying…).

2.1.4.3.3 NHE: Nielsen's Heuristic Evaluation (Nielsen, 1993a)

Heuristic evaluation is a usability engineering method which tries to identify usability problems in a user interface. The identified problems are than used as inputs to the iterative design process. The method requires the availability of a set of evaluators who will test drive the interface. The evaluators are assisted with an observer while they are using the interface. Evaluators can ask questions and request help from the observer. This is the in contrast to the general user testing methods. The evaluators state each recognized usability problem in a session with a reference to the specific heuristic. Therefore it is not enough for an evaluator to just say that "I did not like this" but also he / she should clearly explain the problem with a reference to the specific heuristic in hand.

Nielsen's method is amongst the most popular usability inquiry methods. Heuristics are list of recognized usability principles. These heuristics are general rules to a usable interface. Heuristic evaluation is an ideal method to deploy in the very early stages of the design process due to the reason that it is possible to conduct heuristic evaluation even on a system which is on paper. This is a clear advantage of this method from the perspective of system designers.

Ten usability heuristics that have been suggested by Nielsen (1994b) can be examined below. These are the revised heuristics originally developed with Rolf Molich in 1990 (Molich, Nielsen 1990). Nielsen further examined 249 usability problems in an interface and applied factor analysis to the data (Nielsen 1994a) and came up with these heuristics. These are applicable to almost any user interface. In Nielsen's own words: "They are more in the nature of rules of thumb than specific usability guidelines".

1. Visibility of system status
   The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.

2. Match between system and the real world

   The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.

3. User control and freedom

   Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.

4. Consistency and standards

   Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.

5. Error prevention

   Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.

6. Recognition rather than recall

   Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.

7. Flexibility and efficiency of use

   Accelerators -- unseen by the novice user -- may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.

8. Aesthetic and minimalist design

   Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.

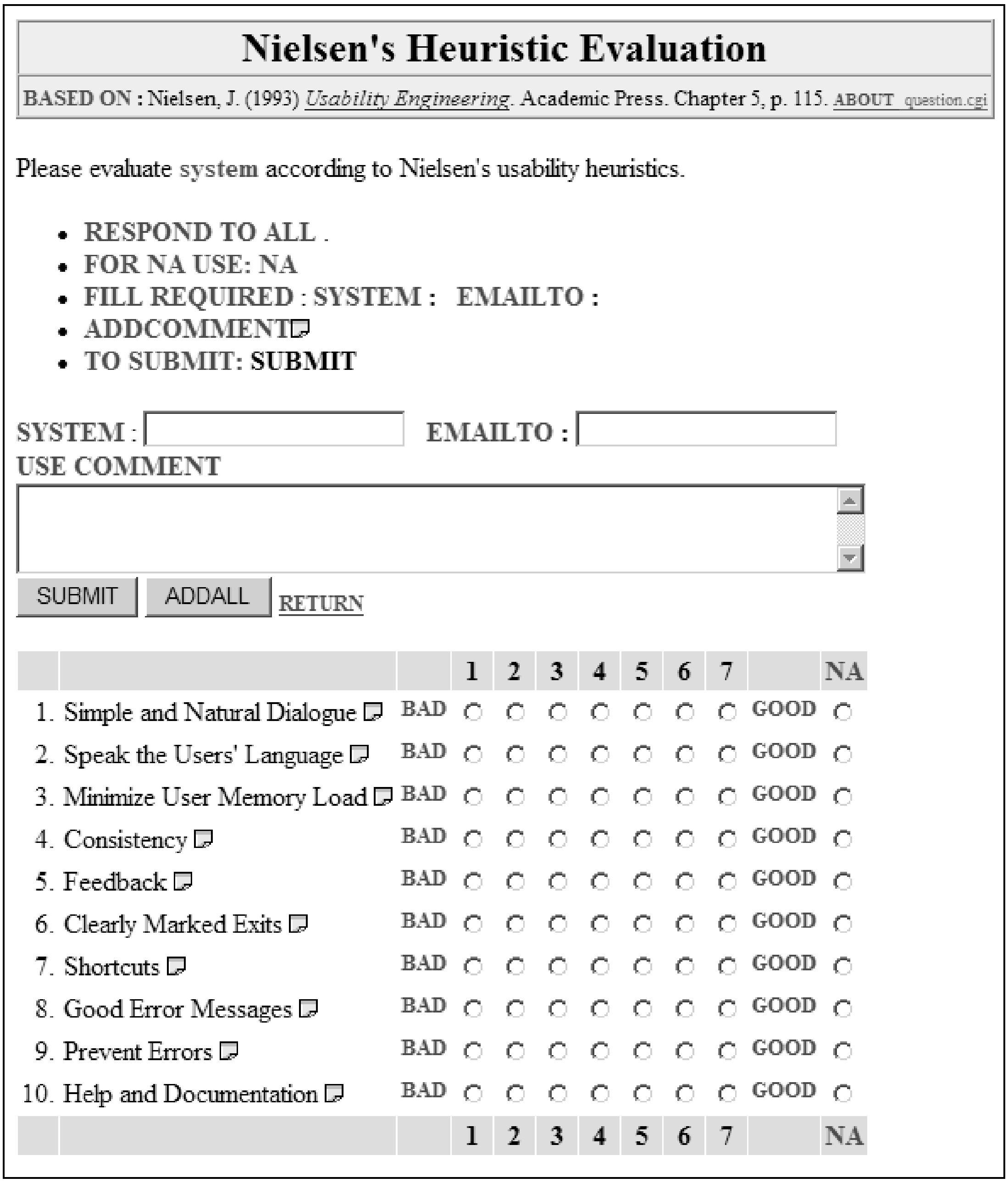
9. Help users recognize, diagnose, and recover from errors

   Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.

10. Help and documentation

    Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

A questionnaire based upon Nielsen’s heuristic evaluation can be examined below in Figure 2.10. The questionnaire is composed of 7-points Lickert scale items.



Nielsen's Heuristic Evaluation

BASED ON : Nielsen, J. (1993) *Usability Engineering*. Academic Press. Chapter 5, p. 115. ABOUT question.cgi

Please evaluate system according to Nielsen's usability heuristics.

- RESPOND TO ALL .
- FOR NA USE: NA
- FILL REQUIRED : SYSTEM : EMAILTO :
- ADDCOMMENT
- TO SUBMIT: SUBMIT

SYSTEM : EMAILTO :

USE COMMENT

SUBMIT ADDALL RETURN

		1	2	3	4	5	6	7		NA
1. Simple and Natural Dialogue	BAD	○	○	○	○	○	○	○	GOOD	○
2. Speak the Users' Language	BAD	○	○	○	○	○	○	○	GOOD	○
3. Minimize User Memory Load	BAD	○	○	○	○	○	○	○	GOOD	○
4. Consistency	BAD	○	○	○	○	○	○	○	GOOD	○
5. Feedback	BAD	○	○	○	○	○	○	○	GOOD	○
6. Clearly Marked Exits	BAD	○	○	○	○	○	○	○	GOOD	○
7. Shortcuts	BAD	○	○	○	○	○	○	○	GOOD	○
8. Good Error Messages	BAD	○	○	○	○	○	○	○	GOOD	○
9. Prevent Errors	BAD	○	○	○	○	○	○	○	GOOD	○
10. Help and Documentation	BAD	○	○	○	○	○	○	○	GOOD	○
		1	2	3	4	5	6	7		NA

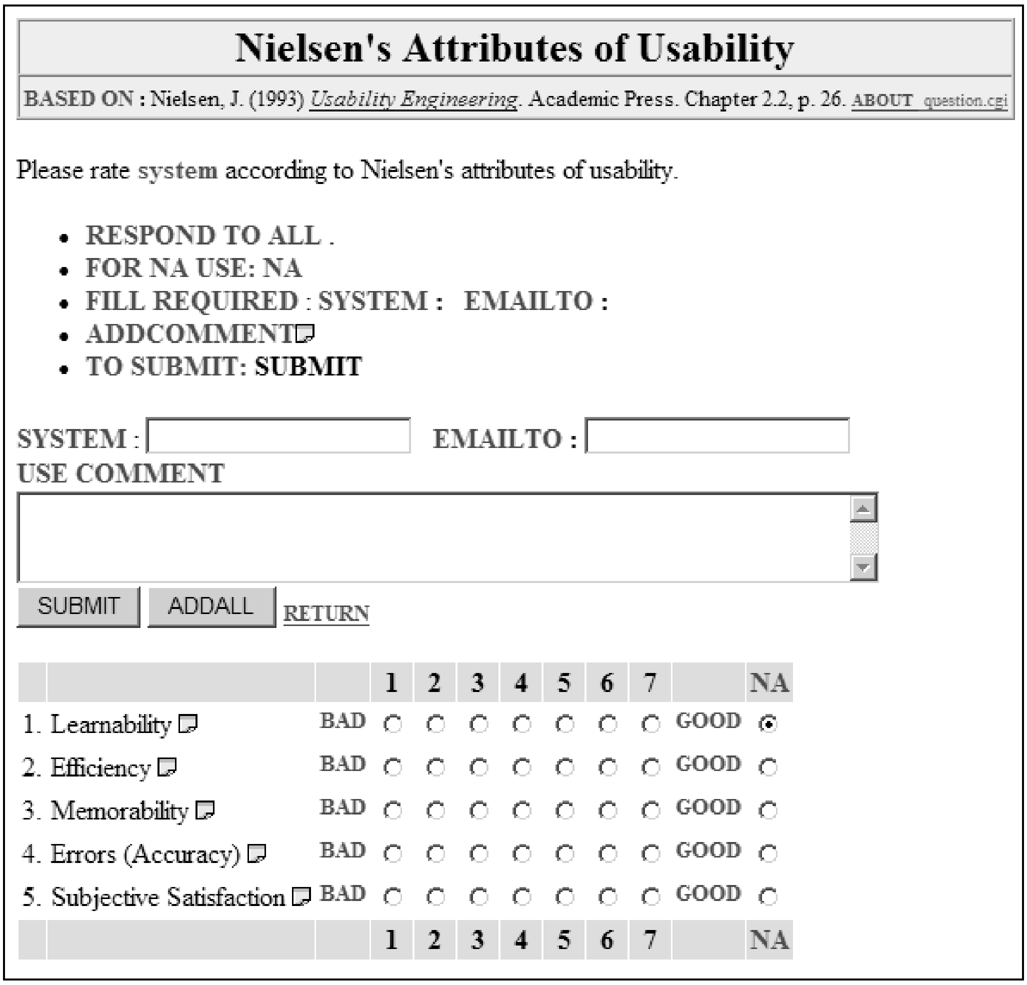
Figure 2.10: Nielsen’s Heuristic Evaluation Questionnaire

Nielsen's method has acquired much publicity and popularity in the field of usability engineering due to the reason that it is relatively cheaper to deploy, quickly to acquire and analyze data, possible to deploy the method very early in the design process thus identifying problems before they are developed and built into the product. It is possible to use the method for both products on the development stage and for finished products. Readily available heuristics and readily available solutions to those heuristics supply extra added value. Therefore as the method is used over and over again by practioners throughout the globe, universal database of common heuristics for a given scenario is built including the fixes to those design problems.

2.1.4.3.4 NAU: Nielsen's Attributes of Usability (Nielsen, 1993a)

As examined before while explaining the different definitions of usability, Nielsen's attributes of usability focuses on the interface usability of a system. One can refer back to the figure in page 5 for more information where usability (interface usability) is a sub-dimension of the usefulness dimension. Nielsen's attributes of usability is a real simple 7 points Lickert scale questionnaire which is composed of 5 items. It is easy and fast to deploy. An example of this questionnaire can be seen in Figure 2.11.

NAU is ideal in cases when one is to examine the relative improvements in a system and the impact of these upon usability. Therefore is a good tool for iterative design process. Nielsen supplies good information regarding the usage of this instrument for measuring the impact of the relative improvements. In the examination of results, he suggests the usage of normalized data for usability dimensions. The initial version of the interface is pointed as "100" and the successor versions results are normalized results to the initial version. Thus an observer of the results can say that if the version 2 has "133" points on the dimension learnability then there is an improvement of 33% over version 1 on version 2. Nielsen also gives information regarding to rank order the potential improvements which will have greater impact over other improvements upon usability. He also supplies information regarding to when to stop iterating through the process as extra iterations marginal utility will be less significant.



Nielsen's Attributes of Usability

BASED ON : Nielsen, J. (1993) *Usability Engineering*. Academic Press. Chapter 2.2, p. 26. ABOUT question.cgi

Please rate system according to Nielsen's attributes of usability.

- RESPOND TO ALL .
- FOR NA USE: NA
- FILL REQUIRED : SYSTEM : EMAILTO :
- ADDCOMMENT
- TO SUBMIT: **SUBMIT**

SYSTEM : EMAILTO :

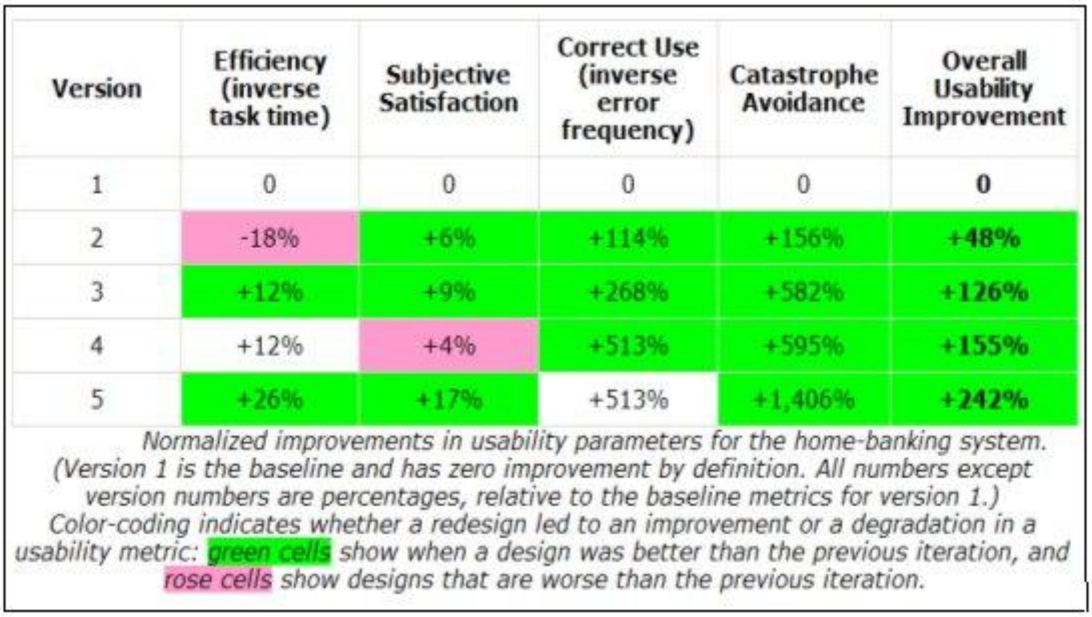
USE COMMENT

SUBMIT ADDALL RETURN

			1	2	3	4	5	6	7		NA
1.	Learnability	BAD	○	○	○	○	○	○	○	GOOD	◉
2.	Efficiency	BAD	○	○	○	○	○	○	○	GOOD	○
3.	Memorability	BAD	○	○	○	○	○	○	○	GOOD	○
4.	Errors (Accuracy)	BAD	○	○	○	○	○	○	○	GOOD	○
5.	Subjective Satisfaction	BAD	○	○	○	○	○	○	○	GOOD	○
			1	2	3	4	5	6	7		NA

Figure 2.11: Nielsen’s Attributes of Usability Questionnaire

The results of NAU study of Danish bank system (Nielsen 1993b) can be examined in Table 2.3. The user interface has been revised 5 times and as sees the relative improvements of the successor designs are measured on 4 scales: efficiency, subjective satisfaction, correct use, catastrophe avoidance.



Version	Efficiency (inverse task time)	Subjective Satisfaction	Correct Use (inverse error frequency)	Catastrophe Avoidance	Overall Usability Improvement
1	0	0	0	0	**0**
2	-18%	+6%	+114%	+156%	**+48%**
3	+12%	+9%	+268%	+582%	**+126%**
4	+12%	+4%	+513%	+595%	**+155%**
5	+26%	+17%	+513%	+1,406%	**+242%**

Normalized improvements in usability parameters for the home-banking system. (Version 1 is the baseline and has zero improvement by definition. All numbers except version numbers are percentages, relative to the baseline metrics for version 1.) Color-coding indicates whether a redesign led to an improvement or a degradation in a usability metric: green cells show when a design was better than the previous iteration, and rose cells show designs that are worse than the previous iteration.

Table 2.3: Improvements in Usability Parameters Measured via Nielsen's Attributes of Usability for a Home Banking System

In this study users were given four sets of tasks:

1. Basic tasks operating on the customer's own accounts
   a. Find out the balance for all your accounts.
   b. Transfer an amount from one of your accounts to the other.
   c. Investigate whether a debit card transaction has been deducted from the account yet.
2. Money transfers to accounts owned by others
   a. Order an electronic funds transfer to pay your March telephone bill.
   b. Set up a series of electronic funds transfers to pay monthly installments for a year on a purchase of a stereo set.
   c. Investigate whether the telephone bill transfer has taken place yet.
3. Foreign exchange and other rare tasks
   a. Order Japanese Yen corresponding to the value of 2,000 Danish Kroner.
   b. Order 100 Dutch Guilders.
   c. Order an additional account statement for your savings account.
4. Special tasks

a. Given the electronic funds transfers you have set up, what will the balance be on August 12th? (Do not calculate this manually; find out through the system).
b. You returned the stereo set to the shop, so cancel the remaining installment payments. ,

Both objective data such as numbers of errors noted by observers and subjective data such as subjective satisfaction is collected. The overall usability improvement score is the geometric mean of the four scales usability improvement scores.

Nielsen's NAU (attributes of usability) is also a great tool amongst his NHE (heuristic evaluation) for being deployed in an iterative design process. NAU is also amongst the very popular definitions of usability in the literature.

2.1.4.3.5 PSSUQ: Post-Study System Usability Questionnaire (Lewis 1992a)

The Post study system usability questionnaire measures 6 system characteristics in a usability context. These characteristics are:

1. Ease of Use
2. Ease of Learning
3. Simplicity
4. Effectiveness
5. Information
6. User Interface

If one refers back to the different definitions of usability, it is quite easy to find out that some of these characteristics are dimensions of the narrow view of usability. Concluding PSSUQ measures mostly the user interface usability of a system.

Post study system usability questionnaire of Lewis consists of 19 items. All items are 7-point Lickert scale. Lewis's original work (psychometric evaluation) on these 19 items found out that the data gathered with this instrument indicated 3 factors (he conducted an exploratory principal factor analysis). These 3 factors explained the 87% of the variance in the data. The items, not considered for any of the factors, were left out due to the reason that they loaded highly on more than 1 scale and thus they were ambiguous. Lewis named these three factors as:

1. System Usefulness (SYSUSE)
2. Information Quality (INFOQUAL)
3. Interface Quality (INTERQUAL).

The internal reliability of the questionnaire in that work is found to be .97. He also examined the validity of the instrument with another usability instrument's finding (ASQ, after study questionnaire) and found that the results are highly correlated (r (20) =.80, p=.0001).

PSSUQ can be observed in Figure 2.12.

1. Overall, I am satisfied with how easy it is to use this system.

STRONGLY AGREE 1 2 3 4 5 6 7 **STRONGLY DISAGREE**

COMMENTS:

2. It was simple to use this system.

STRONGLY AGREE 1 2 3 4 5 6 7 **STRONGLY DISAGREE**

COMMENTS:

3. I could effectively complete the tasks and scenarios using this system.

STRONGLY AGREE 1 2 3 4 5 6 7 **STRONGLY DISAGREE**

COMMENTS:

4. I was able to complete the tasks and scenarios quickly using this system.

STRONGLY AGREE 1 2 3 4 5 6 7 **STRONGLY DISAGREE**

COMMENTS:

5. I was able to efficiently complete the tasks and scenarios using this system.

STRONGLY AGREE 1 2 3 4 5 6 7 **STRONGLY DISAGREE**

COMMENTS:

6. I felt comfortable using this system.

STRONGLY AGREE 1 2 3 4 5 6 7 **STRONGLY DISAGREE**

COMMENTS:

7. It was easy to learn to use this system.

STRONGLY AGREE 1 2 3 4 5 6 7 **STRONGLY DISAGREE**

COMMENTS:

8. I believe I could become productive quickly using this system.

STRONGLY AGREE 1 2 3 4 5 6 7 **STRONGLY DISAGREE**

COMMENTS:

Figure 2.12: PSSUQ Questionnaire Example

9. The system gave error messages that clearly told me how to fix problems.

STRONGLY AGREE	1	2	3	4	5	6	7	STRONGLY DISAGREE

COMMENTS:

10. Whenever I made a mistake using the system, I could recover easily and quickly.

STRONGLY AGREE	1	2	3	4	5	6	7	STRONGLY DISAGREE

COMMENTS:

11. The information (such as on-line help, on-screen messages and other documentation) provided with this system was clear.

STRONGLY AGREE	1	2	3	4	5	6	7	STRONGLY DISAGREE

COMMENTS:

12. It was easy to find the information I needed.

STRONGLY AGREE	1	2	3	4	5	6	7	STRONGLY DISAGREE

COMMENTS:

13. The information provided for the system was easy to understand.

STRONGLY AGREE	1	2	3	4	5	6	7	STRONGLY DISAGREE

COMMENTS:

14. The information was effective in helping me complete the tasks and scenarios.

STRONGLY AGREE	1	2	3	4	5	6	7	STRONGLY DISAGREE

COMMENTS:

15. The organization of information on the system screens was clear.

STRONGLY AGREE	1	2	3	4	5	6	7	STRONGLY DISAGREE

COMMENTS:

Note: *The interface includes those items that you use to interact with the system. For example, some components of the interface are the keyboard, the mouse, the screens (including their use of graphics and language).*

Figure 2.12: PSSUQ Questionnaire Example (contin'd)

16. The interface of this system was pleasant.

STRONGLY AGREE	1	2	3	4	5	6	7	STRONGLY DISAGREE

COMMENTS:

17. I liked using the interface of this system.

STRONGLY AGREE	1	2	3	4	5	6	7	STRONGLY DISAGREE

COMMENTS:

18. This system has all the functions and capabilities I expect it to have.

STRONGLY AGREE	1	2	3	4	5	6	7	STRONGLY DISAGREE

COMMENTS:

19. Overall, I am satisfied with this system.

STRONGLY AGREE	1	2	3	4	5	6	7	STRONGLY DISAGREE

COMMENTS:

Figure 2.12: PSSUQ Questionnaire Example (contin'd)

First 8 items of the questionnaire measures system use (items 1-8), the following 6 items measures information quality (items 9-15) and following 3 items (items 16-18) measures interface quality. Overall usability of the system is measured by averaging the results of all items.

2.1.4.3.6 CSSUQ: Computer System Usability Questionnaire (Lewis 1992b)

Computer System Usability Questionnaire (19 items) is identical to PSSUQ (19 items). The only difference is the wording of the items. In CSSUQ the wordings does not refer to a usability testing situation. An example will better explain this fact:

Item 3 of PSSUQ:

"I could effectively complete the tasks and scenarios using this system"

Item 3 of CSSUQ:

"I can effectively complete my work using this system."

After developing PSSUQ, Lewis was not satisfied with the relatively small sample size of the study. The original PSSUQ was psychometrically analyzed from data gathered from 48 participants. He was also suspicious regarding to the influence of the laboratory setting which the PSSUQ was conducted in. Therefore he revised PSSUQ and developed CSSUQ and he collected data from a mail survey of 825 IBM employees that are spread around the United States.

> The PSSUQ research was preliminary for two reasons. First, the sample size for the factor analysis was small, consisting of data from only 48 participants. Second, the PSSUQ data came from a usability study. This setting may have influenced the correlations among the items and, therefore, the resultant factors. The purpose of this research (Lewis, 1992a) was to use a slightly revised version of the PSSUQ, the computer System Usability Questionnaire (CSUQ) to obtain a database of sufficient size to calculate stable factors from a mailed survey. If the same factors emerged from this research as from the PSSUQ research, the study would demonstrate the potential usefulness of the questionnaire across different user groups and different research settings. (Lewis 1992a)

377 employees returned completed questionnaires. The principal factor analyses results were identical to PSSUQ findings. A three factor solution, as in the case of PSSUQ, was evident. Two of the items in PSSUQ which were found to be ambiguous regarding to which factor they should belong to was also solved. Item 8 of PSSUQ was found to be an item of Factor 1 and Item 15 was found to be a factor of Factor 2. With these results the three factors explained the 98.6% of the variance in the data which is a significant improvement over 87% of PSSUQ. The coefficient of alpha was .95.

The findings of the CSSUQ confirmed that the three factors in PSSUQ were consistent. Both of these instruments are good instruments for after scenario testing purposes.

2.1.4.3.7 ASQ: After scenario questionnaire (Lewis 1995)

After Scenario questionnaire is a three-item questionnaire that IBM uses to assess participants' satisfaction of a specific system usage. The items are 7-point Lickert scale. As the number of items is small it is very easy to apply this questionnaire. It measures three aspects of user satisfaction; ease of task completion, time to complete a task, and adequacy of support information (on-line help, messages, and documentation).

Lewis has psychometrically evaluated ASQ in 48 participant study. The participants has conducted 8 different scenarios and completed the ASQ questionnaires afterwards. An observer logged the activities of participants whether they succeeded or failed the scenario.

After the study he conducted a factor analysis (analysis by scenarios). Examination of 8 factors (with 4 items under each factor) resulted in a coefficient of alpha which exceeded .90. The correlation between the ASQ scores and the logs of the observers (whether the participant failed or succeeded) was -.40 ($p<.01$) meaning participants who successfully completed the scenario gave more favorable ratings. This was an evidence of concurrent validity. The ASQ can be seen in Figure 2.13.

The IBM Questionnaires

The After-Scenario Questionnaire (ASQ)

Administration and Scoring. Give the questionnaire to a participant after he or she has completed a scenario during a usability evaluation. Average (with the arithmetic mean) the scores from the three items to obtain the ASQ score for a participant's satisfaction with the system for a given scenario. Low scores are better than high scores due to the anchors used in the 7-point scales. If a participant does not answer an item or marks N/A, average the remaining items to obtain the ASQ score.

Instructions and Items. The questionnaire's instructions and items are:

For each of the statements below, circle the rating of your choice.

1. Overall, I am satisfied with the ease of completing this task.

   STRONGLY AGREE 1 2 3 4 5 6 7 STRONGLY DISAGREE

2. Overall, I am satisfied with the amount of time it took to complete this task.

   STRONGLY AGREE 1 2 3 4 5 6 7 STRONGLY DISAGREE

3. Overall, I am satisfied with the support information (on-line help, messages, documentation) when completing this task.

   STRONGLY AGREE 1 2 3 4 5 6 7 STRONGLY DISAGREE

Figure 2.13: ASQ Questionnaire Example

2.1.4.3.8 SUMI: Software Usability Measurement Inventory (HFRG)

Software Usability Measurement Inventory is a questionnaire which aims to measure the software quality from the end user's point of view. J. Kirakowski has leaded the team which developed SUMI. It is a method which can be used to assess a full functional product or a prototype. It is one of the few questionnaires which have been roughly examined and its reliability tested amongst several different samples.

Therefore examination of SUMI and its development stage will be supplied in details.

The original work on SUMI has started on 1990. The objectives were*:

1. To examine the CUSI Competence scale and to expand it and to extract further subscales if warranted by the evidence;
2. To achieve an international standardization database for the new questionnaire and to validate its use in commercial environments.

The team that leaded SUMI development project started with an item pool of over 150 items. These items were assembled from previous studies (some mentioned above), discussions with actual end users and from suggestions by HCI (Human Computer Interface) experts. After potential item selection, the subject matter experts classified these items under their selected categories by perceived meaning. After that, some items were rewritten or eliminated if they produced inconsistent allocations in classifications. Kirakowski says that they have chosen Lickert scaling approach because this is considered to be a natural way of eliciting opinions about a software product. We have seen this numerous times in the popular questionnaires that have been examined above. They kept the number of items under each specific category relatively high to overcome variability due to extraneous or irrelevant factors. All these preliminary work resulted in the first draft of SUMI which consisted of 75 items. The Lickert scale used was 5-points Lickert scale with strongly agree in one side and strongly disagree in the other. The neutral point was "Didn't know".

They submitted the first draft to 139 end users from a varying number of organizations. Each user evaluated the software which they were using in their

* HFRG, Human Factors Research Group SUMI website. Available from http://sumi.ucc.ie/sumipapp.html#sumidev , 04-01-2008

organization (thus no single software evaluated). A factor analysis was conducted and five to six different groups formed which gave acceptable internal consistency. The questionnaire was revised and the number of factors was set two five. The ten items which loaded most with the subsequent factors were selected for each factor thus resulting in a draft 2 version of SUMI which consisted of 50 items. The 5-point Lickert scale was also modified and a 3-point scale was used. The response categories were “Agree”, “Don’t know” and “Disagree”.

The revised SUMI was submitted for the second-time to a new sample of 143 users in a commercial environment. Analysis of the results gathered from these users showed that scale reliabilities (5 scales) and item-scale correlations were similar or even better than the results gathered from the original 139 user first sample. ANOVA tests successfully differentiated between different software systems that were evaluated by the users in the sample. Some items are further modified after these results and the selected 5 scales were labeled as;

1. Efficiency: measures the degree to which users feel that the software assists them in their work and is related to the concept of transparency
2. Affect (this scale was also existent in the original work of HFRG’s work on CUSI): measures the user's general emotional reaction to the software -- it may be glossed as Likeability
3. Helpfulness: measures the degree to which the software is self-explanatory, as well as more specific things like the adequacy of help facilities and documentation
4. Control: measures the extent to which the user feels in control of the software, as opposed to being controlled by the software, when carrying out the task
5. Learnability: measures the speed and facility with which the user feels that they have been able to master the system, or to learn how to use new features when necessary.

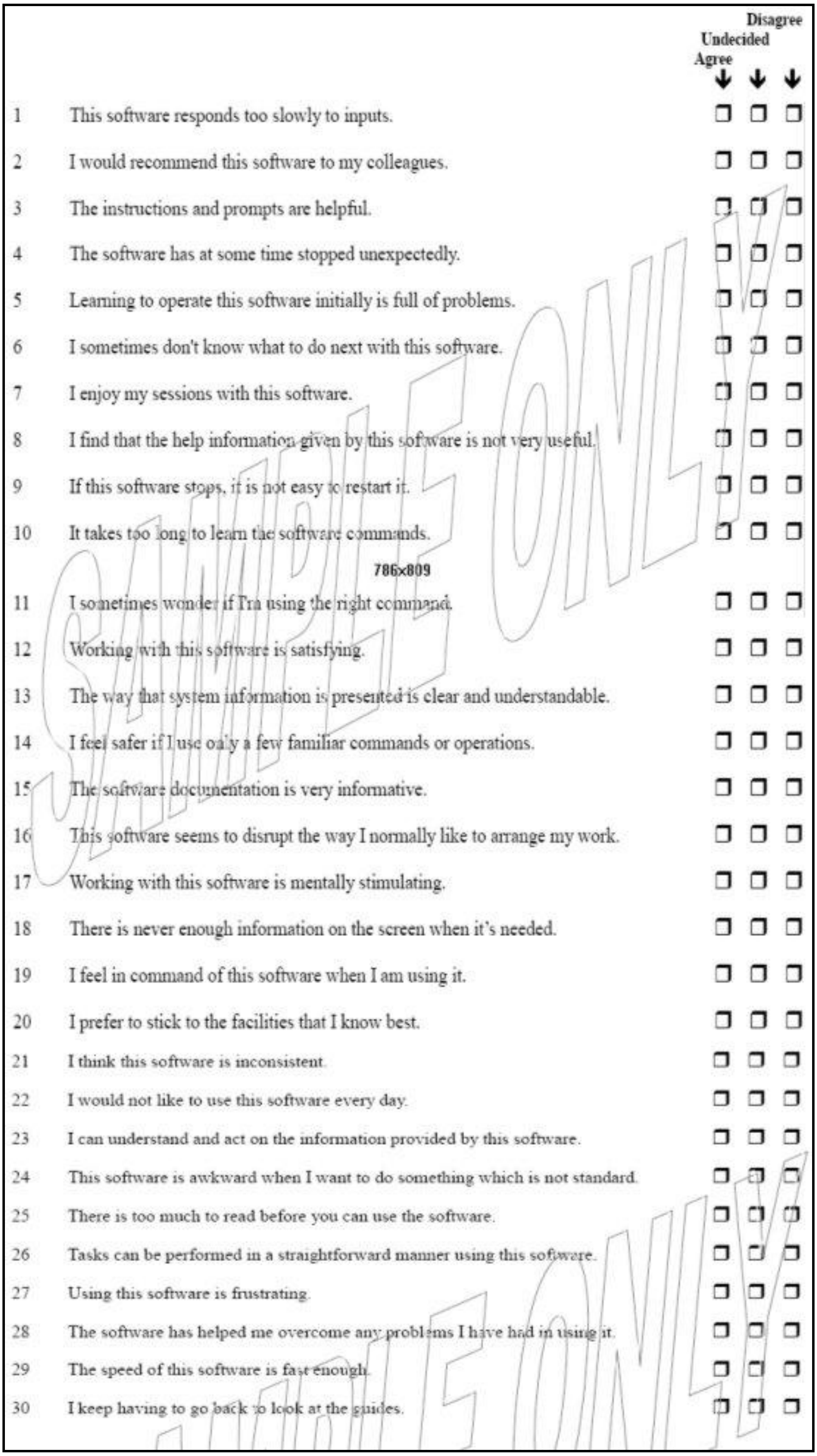
After the second field study of the development team, HFRG group evaluated SUMI in a single organization's context. In this organization the programmer's team was divided into two distinct groups using different versions of the same software which programmer's use daily for building commercial applications). Version 2 of the software was highly popular over version 1 of the same software. This was a know fact amongst the developers and was a common feedback to the manager of the subsequent department in interest. SUMI successfully differentiated between the two versions of the software. In 4 of the 5 scales (only learnability was similar amongst the two versions) version 2 outperformed version 1 according to SUMI results. This has also proved that SUMI was a valid instrument.

The HFRG later selected 25 items out of the reliable and valid 50 items questionnaire and named the new derivative as "Global Scale". The field study for this scale was conducted in a sample of 1100 participants who evaluated well over 150 different systems. The reliability results when compared with the 143 sample field study were consistent, with the global scale being relatively more reliable, amongst the 5 different dimensions. HFRG group examined the underlying factors by deploying Spearman's rho and according to this evaluation control and efficiency factors were found out to be closely correlated. Kirakowski pointed out that future works on SUMI, should examine the control dimension in greater detail.

The final evaluation of the reliability of SUMI was also applied in 1994 and results were consistent with the 1100 participant study which was conducted before.

With this much of psychometric evaluation, SUMI has earned great popularity. It has proven to be one of the most popular and trusted tool which is to be used for evaluation of end user's perception of software quality.

SUMI (the latest and up-to date revision was published in 1993) is composed of 50 statements. The user has to respond to these statements that they "Agree", "Don't Know" or "Disagree". Figures 2.14 demonstrate the 50 statements of SUMI.



		Agree	Undecided	Disagree
1	This software responds too slowly to inputs.	❐	❐	❐
2	I would recommend this software to my colleagues.	❐	❐	❐
3	The instructions and prompts are helpful.	❐	❐	❐
4	The software has at some time stopped unexpectedly.	❐	❐	❐
5	Learning to operate this software initially is full of problems.	❐	❐	❐
6	I sometimes don't know what to do next with this software.	❐	❐	❐
7	I enjoy my sessions with this software.	❐	❐	❐
8	I find that the help information given by this software is not very useful.	❐	❐	❐
9	If this software stops, it is not easy to restart it.	❐	❐	❐
10	It takes too long to learn the software commands.	❐	❐	❐
11	I sometimes wonder if I'm using the right command.	❐	❐	❐
12	Working with this software is satisfying.	❐	❐	❐
13	The way that system information is presented is clear and understandable.	❐	❐	❐
14	I feel safer if I use only a few familiar commands or operations.	❐	❐	❐
15	The software documentation is very informative.	❐	❐	❐
16	This software seems to disrupt the way I normally like to arrange my work.	❐	❐	❐
17	Working with this software is mentally stimulating.	❐	❐	❐
18	There is never enough information on the screen when it's needed.	❐	❐	❐
19	I feel in command of this software when I am using it.	❐	❐	❐
20	I prefer to stick to the facilities that I know best.	❐	❐	❐
21	I think this software is inconsistent.	❐	❐	❐
22	I would not like to use this software every day.	❐	❐	❐
23	I can understand and act on the information provided by this software.	❐	❐	❐
24	This software is awkward when I want to do something which is not standard.	❐	❐	❐
25	There is too much to read before you can use the software.	❐	❐	❐
26	Tasks can be performed in a straightforward manner using this software.	❐	❐	❐
27	Using this software is frustrating.	❐	❐	❐
28	The software has helped me overcome any problems I have had in using it.	❐	❐	❐
29	The speed of this software is fast enough.	❐	❐	❐
30	I keep having to go back to look at the guides.	❐	❐	❐

Figure 2.14: SUMI Questionnaire Example

		Agree	Undecided	Disagree
31	It is obvious that user needs have been fully taken into consideration.	❐	❐	❐
32	There have been times in using this software when I have felt quite tense.	❐	❐	❐
33	The organisation of the menus or information lists seems quite logical.	❐	❐	❐
34	The software allows the user to be economic of keystrokes.	❐	❐	❐
35	Learning how to use new functions is difficult.	❐	❐	❐
36	There are too many steps required to get something to work.	❐	❐	❐
37	I think this software has made me have a headache on occasion.	❐	❐	❐
38	Error prevention messages are not adequate.	❐	❐	❐
39	It is easy to make the software do exactly what you want.	❐	❐	❐
40	I will never learn to use all that is offered in this software.	❐	❐	❐
41	The software hasn't always done what I was expecting.	❐	❐	❐
42	The software has a very attractive presentation.	❐	❐	❐
43	Either the amount or quality of the help information varies across the system.	❐	❐	❐
44	It is relatively easy to move from one part of a task to another.	❐	❐	❐
45	It is easy to forget how to do things with this software.	❐	❐	❐
46	This software occasionally behaves in a way which can't be understood.	❐	❐	❐
47	This software is really very awkward.	❐	❐	❐
48	It is easy to see at a glance what the options are at each stage.	❐	❐	❐
49	Getting data files in and out of the system is not easy.	❐	❐	❐
50	I have to look for assistance most times when I use this software	❐	❐	❐

Please check you have ticked each item.
Thank you.

Figure 2.14: SUMI Questionnaire Example (contin'd)

SUMI's objective is not only to measure current usability status of a software package but also diagnose usability problems. By this way, the tool is of value to system designers too. As seen in Nielsen's heuristic approach, for an instrument to have value, not only by end users, it also has to diagnose and pinpoint the problems which will impact the perceived usability levels to a significant degree. HFRG developed the "Item Consensual Analysis" feature for this purpose. This is accomplished by examining the observed patterns (how many Agrees, Don't Agree, Don't Knows) with the expected patterns of the standardization database (this database has been formed while SUMI was being developed for the purpose of creating a formula which will standardize responses to each specific item of SUMI. With this standardization each scale has got a mean of 50 with a standard deviation of 10.) . Items which have large discrepancy between the expected and observed patterns of responses are the ones which are unique to the system under observation. These unique responses can elicit positive or negative comments.

2.1.4.3.9 MUMMS: Measurement of Usability of Multimedia Software (HFRG)

Measurement of usability of multimedia software questionnaire, as SUMI, was also developed by the Human Factors Research Group at University College, Cork. The reason to develop a new questionnaire, when already having a very successful and welcomed questionnaire is explained by HFRG's own statement below:

> In the early 1990s, SUMI swept the board, so that by now every major multinational IT company uses SUMI in some part of its enterprise, and many consultancies and SMEs in Europe and the United States also use it. The time has come for a major re-think of end-user based questionnaires, and the MUMMS questionnaire is being developed in response to the rapidly changing patterns and technology of computing today. Multi-media computer products are establishing themselves as part of the market and it is becoming necessary to develop ways in which these products can be assessed for quality of use by the end users themselves. [*]

The dimensions which emerged from MUMMS development work has been the same as SUMI; affect (aka likeability), control, efficiency, helpfulness and learnability. Interestingly the HFRG group, when analyzing (factor analysis) the data

[*] HFRG (MUMMS), Human Factors Research Group MUMMS website. Available from http://www.ucc.ie/hfrg/questionnaires/mumms/info.html , 08-01-2008

gathered from early works on MUMMS, found out a new potential dimension which they prefer to call “Excitement”. The group gives the following definition for this new dimension; “Excitement is the extent to which end users feel that they are 'drawn into' the world of the multi-media application and it seems to capture some of the fascination which the best multi-media apps exercise over their users.”

In a work at Berlin States Museum, the HFRG group evaluated a special software which is developed to present Museum’s cultural assets. The group revised SUMI items for multimedia applications. The resulting questionnaire was also containing 50 items as SUMI. The analyzed data came up with the very same five dimensions that were found in the development of SUMI.

2.1.4.3.10 WAMMI: Website Analysis and Measurement Inventory

Website analysis and measurement inventory was developed by Nigel Claridge and Jurek Kirakowski in 1996. The developers explain its objectives as;

> WAMMI helps you accomplish your business goals in the following ways;
> - Measures user satisfaction of you web site based on user-reactions.
> - Generates objective data for management in easy-to-understand format.
> - Provides a powerful basis for web site changes and design improvements.
> - Benchmarks your site relative to others in terms of user-satisfaction.
> - Tracks web site performance to see if business goals are being accomplished.*

WAMMI is a 20 item questionnaire. The dimensions that are the basis for this questionnaire are a slight modification of the SUMI’s. Controllability, Efficiency, Helpfulness, Learnability are existent however affect (aka likeability) does not exist. The other dimension which is existent is attractiveness.

* WAMMI website, Available from http://www.wammi.com/about.html, 12-01-2008

The developers claim the report to have a reliability of .90 to .93 (Cronbach alpha). As in the case of SUMI and MUMMS, the developers have established a standardization database which contains 200 site evaluations. A sample of WAMMI report can be seen in Figure 2.15.

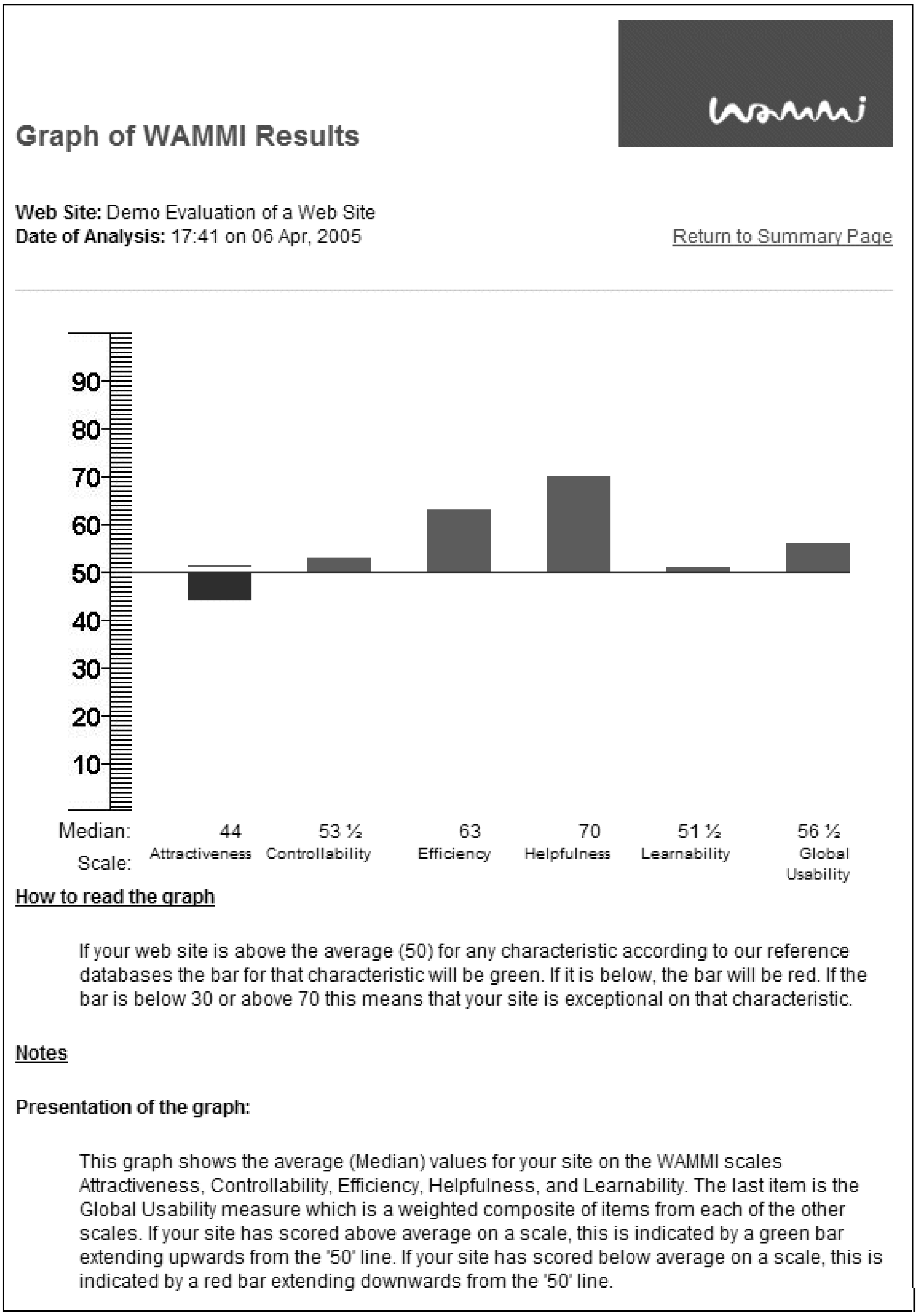


Figure 2.15: Graph of WAMMI Results

2.1.4.3.11 EUCSI: End User Satisfaction Instrument (Doll, 1994)

Doll and Torkzadeh present a model of satisfaction, the End-User Computing Satsifaction Instrument (EUCSI). This instrument is context dependent; its objective is to examine a specific application. The developers define satisfaction of a computer system as “an affective attitude towards a specific computer application by someone who interacts with the application directly”. Reported Cronbach alpha, for EUCSI, range from 0.65 to 0.89. The dimensions of the concept include the following; content, accuracy, format, timeliness and ease of use.

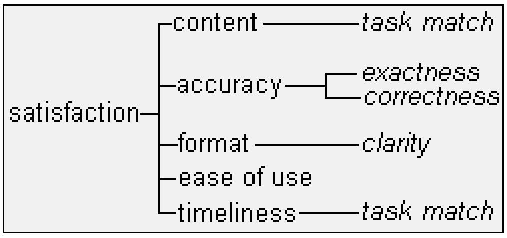


Figure 2.16: End User Satisfaction Model by Doll

1. The 'content' of the system refers to the user's belief concerning the system's ability to provide the information that is needed. The items reflect a passive view of the user in which a computer system provides information without the user's active involvement. The scale makes no reference to users' possibilities of influencing and mastering the tasks they need to accomplish.
2. The concept of 'accuracy' is not specified. Possible references include the precision or correctness of the system. Both are related more to technical reliability than to usability. Reliability is, of course, a relevant criterion from the point of view of user satisfaction, but for most contemporary systems precision and correctness of the technical bases of the systems are taken for granted in user-interface-related approaches.
3. 'Format' refers to beliefs concerning information presentation, the usefulness of the format and clarity.
4. 'Timeliness' seems not really to address the rate of operation or response times, but rather the topicality of information. This, like correctness, seems to

be a suitable criterion for only some specific applications, for instance process control. In most office applications the information is produced by the users themselves through active interaction with the system instead of waiting for the system to provide the information.

5. 'Ease of use' applies expressions such as 'user friendly' and 'easy to use'.

The writer’s did not specify all dimensions of the model precisely such as accuracy. Research done regarding to this instrument’s reliability points has varying degrees of reliability. When compared the dimensions, there are few items explaining usability in the sense used by ISO9126, ISO9241, Nielsen and etc.

2.2 Literature Review Regarding Computer Attitude

In this part of the thesis, an examination of the concept computer attitude is supplied. Its deconstruction into its sub-dimensions and the research done regarding these individual sub-constructs is also analyzed. Finally questionnaires that are developed to measure the concept are introduced.

2.2.1 Introduction to Computer Attitude

Attitudes towards objects is everywhere in life. Every individual has different attitudes towards different objects. Some (attitude) are positive behaviors towards the objects of interest which motivates the individual to use / deploy that object and some are negative which repels the individual to utilize the object. Actually attitude is a very complex behavior which determines, to a very significant degree, ones willingness to test / use / utilize a specific object for a specific purpose. As will be seen in the following sub-sections, it is a real complex construct which is driven by several sub-constructs.

In IT (Information Technology) industry attitude is an important construct which is in-depth examined by researchers in the field. There are several projects which have suffered from employees resisting to utilize the project due to negative attitudes to IT technology and in contrast there are several projects that have been very positively welcomed by employees. The IT professional's major concern is to understand the different behavioral responses of its potential customers towards an IT solution (software, hardware, etc,.) and try to prescribe a recipe for overcoming the attitude barriers.

Important work that is carried out by Fishbein and Ajzen (1975) proposes that attitudes toward an object play an important role in influencing subsequent behaviors. The literature has much focused on this aspect in educational context where technology is being deployed to an increasing amount each and every day. Liaw (2002) explored the impact of attitude towards the acceptance of information technology in an educational context and concluded that if users do not have positive attitudes towards a solution, then they will not use it does not matter how capable it is. As evident, computer attitude is the most important construct of interest to IT industry. Being able to deploy solutions which will be positively welcomed by potential users is of great interest. Although this topic and construct is examined in-depth there is still no universally agreed definition of computer attitude. Kay (1993) located 14 different definitions of the construct. Smith, Caputi and Rawstorne (2000) found out that there is still no universally accepted definition of a construct for computer attitude.

Loyd and Gressard (1984) commented their findings as a general attitude toward working with computers was a reflection of ratings on three factors namely "liking, confidence and freedom from anxiety". Computer experience is a construct which is of great interest to researchers. Generally speaking, positive correlations between computer experience and computer attitude is found. As a general rule, the more contact people have with computers, the more likely they are to express favorable attitudes (Bozionelos 2001). This very fact changes in the case of computer anxious users. Some studies expressed that if an individual has negative reactions to computers, his / her reactions will be reinforced with increased computer experience (Rosen, Sears and Weil 1987).

The researchers in the field tried to differentiate between objective computer experience and subjective computer experience (this will be examined in details later). The distinction between computer attitude and subjective computer experience is somewhat blurred. One can quite easily comment that subjective computer experience (having a good experience while interacting with computers or a negative one and as such the individual tries to run away from interacting with computers) and computer attitude is almost the same thing. However when we focus onto this matter, we can find out individuals who although had prior negative experiences with computers, still do not possess negative attitudes to computers. As evident, there is still much to research on this matter to conclude that; subjective computer experience and computer attitude are the same thing or they are different constructs or subjective computer experience is a dimension of the computer attitude construct.

Existing literature shows that computer confidence is an important variable of interest. Some researchers evaluated this variable as a dimension of the computer attitude construct (Levine & Donitsa-Schmidt (1998), Loyd & Gressard (1984), Shashaani (1994)) whereas some evaluated it as a separate but related construct (Bear, Richards & Lancaster (1987), Gardner, Dukes & Discenza (1993)).

Gender is another variable that has been studied in computer attitude literature. Detailed information regarding gender and its correlation with attitude construct will be supplied in the following sections.

Summing up, important variables that has been studied in the literature regarding computer attitude are; computer experience (objective and subjective), computer anxiety and computer confidence (computer self-efficacy). Figure 2.17 is a sum up of computer attitude and its related constructs.

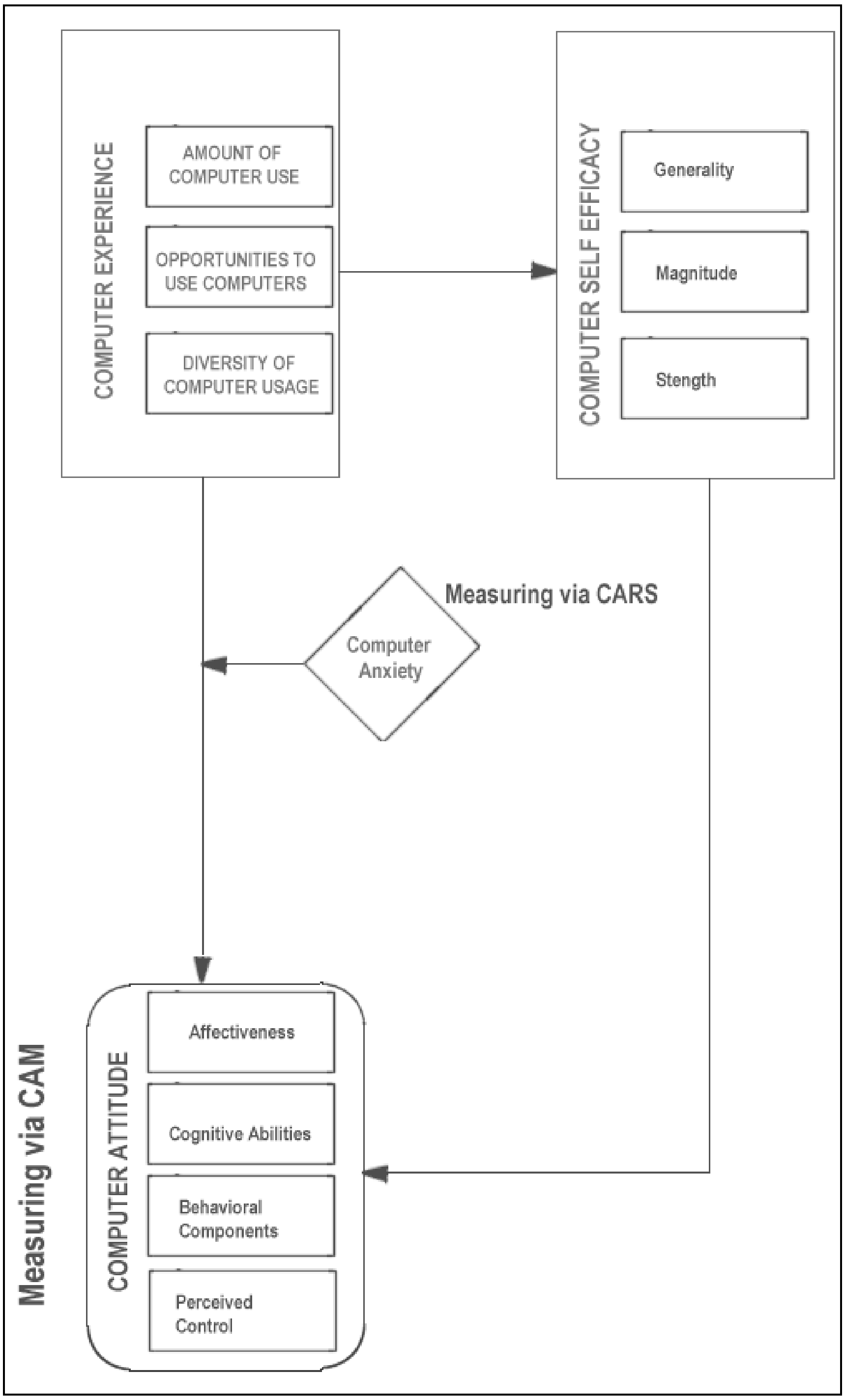


Figure 2.17: General Representation of the Literature Review for Computer Attitude and Its Dimensions

2.2.2 Variables that Has Been Studied Regarding Computer Attitude

Although there is no consensus regarding to the construct, still some variables are almost existent in any of the different definitions and decompositions of computer attitude. These are; computer experience, computer self-efficacy, computer anxiety and demographic variables (especially gender). Each of these variables is being examined in the following sub-sections.

2.2.2.1 Computer Experience

As in the case of computer attitude; there is little agreement concerning the definition of computer experience (Smith, Caputi, Crittenden, Jayasuriya, & Rawstorne, 1999). Most researchers in the field measured computer experience quantifiably, that is, in terms of amount of computer use (such as; how many hours a participant is using spreadsheets in a week). They discriminated between different types of applications (such as spreadsheets, database applications, internet browsing, mail, etc.) and tried to find out a specific usage pattern of applications which will explain computer experience and its relation to computer attitude to a better degree. This approach to computer experience actually measures “computer use” and this resulted the use of both of these terms interchangeably. Smith argued against this indiscriminate use of the terms (Smith, Caputi, Crittenden, Jayasuriya, & Rawstorne, 1999). Computer use and computer experience are two different things. Measuring computer experience quantifiably (such as the amounts of hours spent for a specific type of usage pattern) measures the objective aspect of computer experience. On the other hand a user that has passed through a negative experience with computers is also a different measure of the concept and this aspect of computer experience is labeled as subjective computer experience. A good example for utilizing both of these two aspects on the concept is an anxious user who despite increasing objective computer experience has negative subjective experiences.

The literature is full of research that has only focused on the objective aspect of computer experience which has only focused on the amount of computer use. It was

Jones and Clark (1995) which has enriched this narrow view to objective experience by examining the construct in three dimensions; amount of computer use, opportunities to use computers and diversity of computer experience. Amount of computer use is the same objective aspect of computer experience discussed above. “Opportunities to use computers” covers issues such as whether the user has access to computers at home, the extent of usage and other related issues. Examples to these questions are; having access to computer at home (yes/no), the extent of usage (rarely/sometimes, often), the frequency of general computer use (never/rarely/often/everyday) (Brosnan & Lee, 1998). Diversity of experience examines a person’s usage of software packages, word processing, spreadsheets, databases, programming, games, computer-assisted-learning and familiarity with computer languages, and development of computerized information systems (Igbaria & Chakrabarti, 1990; Jones & Clark, 1995). The responses to these items are usually on 7-points Lickert scales (not at all/less than once a week/about once a week/2 or 3 times a week/4 to 6 times a week/at least once a day/more than once a day).

In contrast to measuring objective computer experience by gauging the user’s usage patterns (hours of usage, etc.) some researchers attempted to measure this concept in terms of skill levels (such as Geissler and Horridge 1993). They ask respondents to rate their skill levels on different tasks, typing, programming, word processing and etc. Supplied scales have 4 or 5 points (poor/fair/good/expert/none to very high).

Other researchers even go further to inquire about perceived computer knowledge for measuring objective computer experience. Computer Understanding and Experience scale (CUE; Potosky & Bobko, 1998) asks respondents about computer terminology, includes questions on computer literacy and etc.

Research done and scales developed are mostly measuring the objective aspect of computer experience. Jones and Clark’s approach is the most detailed one. When one is to become curious about measuring subjective computer experience, it should be noted that there is not much research existent. It should be interesting to examine the impact of the variable “subjective experience” upon computer attitude. The expected correlation can be positive, meaning; users who have good subjective experiences with computers tend to have better computer attitudes.

An example of a questionnaire which tries to measure computer experience, according to the Jones and Clark’s approach, can be seen in Figure 2.18 (Garland, Noyes 2004). The questionnaire tries to measure amount of computer usage (items 1, 2, 3), opportunities to use computers (items 2 (at home), 3 (at university)) and diversity of computer usage (all remaining items).

Relevant sections of the questionnaire

1. For how many years have you used a computer?
2. For how many hours in a week do you use a computer at home?
3. For how many hours in a week do you use a computer at university?
4. Please complete the following:

	How many hours per week do you spend on each of these?	For how many years have you used each of these?	How would you rate your level of experience from 0 = ‘none’, 1 = ‘novice’, 2 = ‘good’, 3 = ‘expert’?
E-mail			
Internet			
Computer games			
Word processing			
Databases			
Spreadsheets			
CAL packages			
Other software			
Programming			

Figure 2.18: An Example of a Questionnaire Which Measures Computer Experience

Throughout the literature there have been several attempts to examine the relationship between computer experience and computer attitude constructs. Most of the instruments developed were shown to be reliable however to a lesser extent valid. This is quite normal as the definition of computer experience even itself is rather ambiguous. Levine and Donitsa-Schmidt (1998) found small to moderate levels of correlations between computer experience and computer attitude. They have measured computer use from three items; whether or not respondents had access to a home computer, the extent of school computer use and frequency of overall use. Dambrot, Watkins-Malek, Silling, Marshall, and Garver (1985) conducted a multiple regression analysts to predict computer attitude. They have measured computer knowledge from whether the participants had completed a course and / or had knowledge of a computing language. Use was measured by asking the respondents their computer usage levels on a 3-points scale (never used, limited use, extensive use). Their work found out that computer experience was a poor predictor of computer attitude. Shashaani (1994) examined four aspects of computer experience, course enrolment, number of computer courses taken, computer ownership and usage (hours each week). Examining each aspect's impact upon computer attitude was the research objective. Hours of computer use each week and number of courses taken showed moderate positive levels of correlations with computer attitude. On the other hand course enrolment and computer ownership showed small positive correlations. Garland and Noyes (2004) commented on Shashaani's work as follows;

> The relationships were largely limited to associations with the utility dimension, which correlated significantly with all use factors. Confidence for females was not significantly correlated with any use measure, which contrasts with all but ownership showing positive, low or moderate correlations for male confidence. The findings suggest that the relationship between computer attitude and experience is not straightforward, but rather that it is dependent on the nature (and quantification) of the experience measure, and the dimensions that are deemed component parts of the attitude construct.

Generally speaking, previous studies have shown that computer experience (Every author defines and explains computer experience dimensions differently to a varying degree) and computer attitude are positively correlated. Some studies show moderate and some show small levels of positive correlations. None of the studies were able to prove that computer experience is a good predictor of computer attitude. The problem resides in computer experience's definition and its selected dimensions. Much research has to be done to come up with a universally welcomed definition of the construct computer experience in order to propose solid dimensions.

2.2.2.2 Computer Self-Efficacy

Computer self efficacy refers to individual confidence in one's capability to use a computer and may help determine ease of skill acquisition (Smith (?)). However higher computer self-efficacy does not always mean higher performance. "In approaching learning tasks, however, those who perceive themselves to be supremely self-efficacious in the undertaking feel little need to invest much preparatory effort in it" stated Bandura (1982). Hackett and Betz examined self-efficacy in math problem solving and found a moderate correlation between math self-efficacy and problem solving (Hackett, Betz 1989). There is much debate on computer self-efficacy in the literature. Levine, and Donitsa-Schmidt (1998) argued that computer self confidence (efficacy) and computer anxiety are essentially the same thing. Beckers and Schmidt (2001) treated computer self efficacy as a dimension of computer anxiety. Coley, Gale, and Harris (1994) treated computer self confidence and computer anxiety as dimensions of the computer attitude.

Computer Self-Efficacy Assessment (CSEA) questionnaire is a 23-item 5-point Lickert scale questionnaire (from completely confident to not at all confident) (Smith (?)). All 23 items can be examined in Table 2.4.

Descriptive Statistics for Computer Self-Efficacy (Perception) and Computer Performance

	Perception				Performance			
Assessment Item	*M*	*SE*	*Min*	*Max*	*M*	*SE*	*Min*	*Max*
Unit 1 - Word Processing								
1. Keyboard documents	3.67	.29	2	5	3.89	.51	1	5
2. Use template	4.11	.26	3	5	3.00	.50	1	5
3. Save and print documents	4.67	.24	3	5	4.67	.33	2	5
4. Edit documents	4.22	.28	3	5	2.67	.50	1	5
5. Ensure quality	3.78	.22	3	5	3.67	.50	1	5
Unit 2 - Spreadsheet								
6. Create spreadsheet	3.89	.26	2	5	3.78	.52	1	5
7. Enhance spreadsheet	4.11	.26	3	5	3.00	.47	1	5
8. Use formulas	3.44	.29	2	5	3.56	.60	1	5
9. Perform special functions	3.44	.29	2	5	3.67	.55	1	5
Unit 3 - Database								
10. Create database	3.00	.29	2	4	3.56	.53	1	5
11. Perform database functions	2.67	.41	1	5	3.44	.38	1	4
Unit 4 - Presentation Graphics								
12. Create visual aids	2.89	.35	2	5	1.78	.46	1	5
13. Incorporate special effects	2.56	.34	1	4	1.89	.48	1	5
Unit 5 - GUI Management								
14. Exchange data between software	2.89	.35	2	5	2.11	.45	1	5
15. Create folders	3.67	.33	2	5	3.00	.65	1	5
16. Use virus detectors	3.22	.43	1	5	3.44	.56	1	5
17. System software commands	3.67	.41	1	5	2.56	.63	1	5
18. Locate hardware devices	3.22	.40	1	5	3.00	.53	1	5
19. Manage files and disks	3.67	.33	2	5	2.44	.58	1	5
Unit 6 - Telecommunications								
20. Communicate using e-mail	4.44	.29	3	5	3.67	.67	1	5
21. Electronic attachments	3.67	.50	1	5	2.22	.62	1	5
22. Electronic communications	3.33	.41	2	5	2.33	.67	1	5
23. Use a web server	4.33	.29	2	5	3.00	.67	1	5

Table 2.4: Descriptive Statistics for Computer Self-Efficacy and Computer Performance from a Sample Questionnaire

There is no consensus regarding to how computer self-efficacy should be treated (whether an independent construct or a dimension of attitude) when examining the relationship of this concept and computer attitude. Still, computer self-efficacy is an important variable of interest when one tries to examine the construct computer attitude.

2.2.2.3 Computer Anxiety

Another important variable of interest that has been studied when trying to explain computer attitude is computer anxiety. There is much research on this variable (Cambre & Cook, 1985, Cohen & Waugh, 1989, Torkzadeh & Angulo, 1992). People who have lower self-efficacy can be more anxious (Brosnan, 1998). Anxious people react with greater magnitudes when something unexpected occurs such as a system error. Brosnan stated that anxious people tend to become even more anxious in the case of an unseen or unknown event occurrence (Brosnan, 1998). On the other hand if one is to increase computer self-efficacy, this will decrease computer anxiety.

> If problems are seen as challenges rather than problems, they may not be as frustrating, which is most likely directly related to level of prior experience as well as computer self-efficacy. This may be due to the perception of locus of control; these individuals understand and can attempt to control the ‘‘problem space’’ they encounter. (Bessie’re, Newhagen, Robinson, Shneiderman 2006)

Computer anxiety can affect performance and attitudes. Increased computer self-efficacy will tend to decrease anxiety. However if one does not increase computer self-efficacy and interact with computers in an increasing manner this can yield to increased anxiety. It is therefore important for companies to assess anxiety levels of their employees before deploying an IT solution. If the existing anxiety levels are found to be high, then first of all increasing computer self-efficacy (by training) has to be the first action. Otherwise increased exposure to the new system can result in such anxiety levels that the deployment of the overall new solution will be much more difficult.

2.2.2.4 Gender

Gender and its relation with computers is a great interest amongst researchers. Much research has been conducted throughout the globe regarding this matter. Generally speaking, most of these researches indicate greater computer experience for males than females. Brosnan & Lee (1998) has found similar results in United Kingdom. Balka & Smith (2000) has found the same symptom in United States. When examined the case for computer attitude, the same stays true for males; there is a tendency to find that females have more negative attitudes towards computers (Durndell & Thomson in UK, 1997, Whitely in USA, 1997). Regarding computer anxiety; there are several researches that point out the same symptom; females are more computer anxious than males (i.e. Maurer, 1994). Research on computer self-efficacy and gender has also found that males on average have more computer self-efficacy than females (Bandura, 1997, Torkzadeh & Koufteros, 1994). Generally speaking in all important variables of computer attitude, males have been found to be more positive in contrast to females. However there are also exceptions. Brosnan & Lee (1998) discovered that in Honk Kong, males were more computer anxious than females. Another interesting exception was found in the ex communist countries. Reinen and Plomp (1997) as a part of a large cross cultural study on school children's knowledge about computers found that the Bulgarian sample in the study provided the smallest gender differences. Even becoming more interesting, Wright (1997) found that in most of majority of the students in math and computing in Bulgaria and Romania were females. Some other studies regarding the ex communist countries found that in some cases females had better computer attitude than males and in some other studies computer self-efficacy was found to be lower in females than males. The paragraph below is a good explanation of the symptom in ex communist countries (Durndell, Haag 2002).

> The particular interest of the emerging ex Communist countries in Eastern Europe is that they were considerably industrialized in the Soviet era. Historically, these previous regimes emphasized both gender equality and the role of technology, and appeared to create a relatively gender neutral technology. Thus, as late as the 1980s, as many if not more females than males were studying to be engineers in these countries, at a time when under 10% were doing so in the UK. It seemed that at the time the sex role expectations in these countries were somewhat different from those of the Western industrialized countries. However, with the collapse of the regimes, the question arose as to whether the apparent relatively gender neutral approach to technology would survive or would turn into the Western pattern of relatively greater male dominance. Would relatively gender neutral attitudes towards technology be deeply embedded or would they be jettisoned as features of the now old discredited Communist society?

Summing up, gender and its relation with computer attitude and its variables is a great research area which has acquired much interest by researchers. Generally speaking, it is possible to indicate that previous research has found that on average males have better computer attitude, are less computer anxious, are more computer self-confident and are spending much more time with computers. In the industrialized world of the western countries, it is quite easy to find that males are dominating the economy and thus such a situation is normal. However when we move to other regimes where gender equality is existent and female participation in areas of specializations which are known to be "man-jobs" are high, there is less difference, between males and females, (and sometimes to the favor of females) in computer attitude and its variables.

2.2.3 Measuring Computer Attitude

Our study mainly does focus on computer attitude and its impact upon perceived usability in a specific context (an e-commerce site, a shopping scenario). Therefore this sub-section of the literature review will yield information only about measuring computer attitude although we have examined such important variables studied regarding the construct such as computer anxiety, computer self-efficacy and computer experience.

There have been several different scales developed throughout the literature for measuring computer attitude. Amongst these scales developed by Loyd and Gressard (1984) and Kay (1989) are popular and psychometrically evaluated ones.

2.2.3.1 Loyd and Gressard (1984) Computer Attitude Measurement Scale

Figure 2.19 displays the items of the scale developed by Loyd and Gressard (1984). The scale has 5 items and all of these items are measured using a 5 point Lickert-scale where 1=strongly disagree and 5=strongly agree.

User attitude

I like working with computers.

I look forward to those aspects of my job that require me to use a computer

Once I start working on the computer, I find it hard to stop

Using a computer is frustrating for me

I get bored quickly when working on a computer

Figure 2.19: Items of Computer Attitude Measurement Scale

2.2.3.2 Kay (1989) CAM Computer Attitude Measurement Scale

Original CAM comprises 10 bipolar dimensions. These are;

- Likeable / Unlikable
- Good / Bad
- Happy / Unhappy
- Comfortable / Uncomfortable
- Calm / Tense
- Full / Empty
- Natural / Artificial
- Exciting / Dull
- Fresh / Suffocating
- Pleasant / Unpleasant

Kay has used 7-point Lickert-scale for each dimension (extremely, moderately, slightly, neither, slightly, moderately, extremely). Some of CAM's dimensions are unclear under specific testing contexts such as natural and artificial. When a user is to evaluate a software solution can one discuss or comment upon whether natural or artificial is better to each other. Further more fresh / suffocating, happy / unhappy dimensions can also be not that suitable for specific contexts. One can refer to Kurosu and Kashimura (1995) for an in-depth evaluation of Kay's CAM scale and its interpretation.

Chapter 3

Theoretical Framework

3.1 Hypothesis Development

In chapter 2, we have examined usability and computer attitude literatures. Examining the relationship between computer attitude and usability is our main objective. Theoretical framework of the study can be examined in Figure 3.1. Hypothesis which will be tested in the research can be found on the next page.

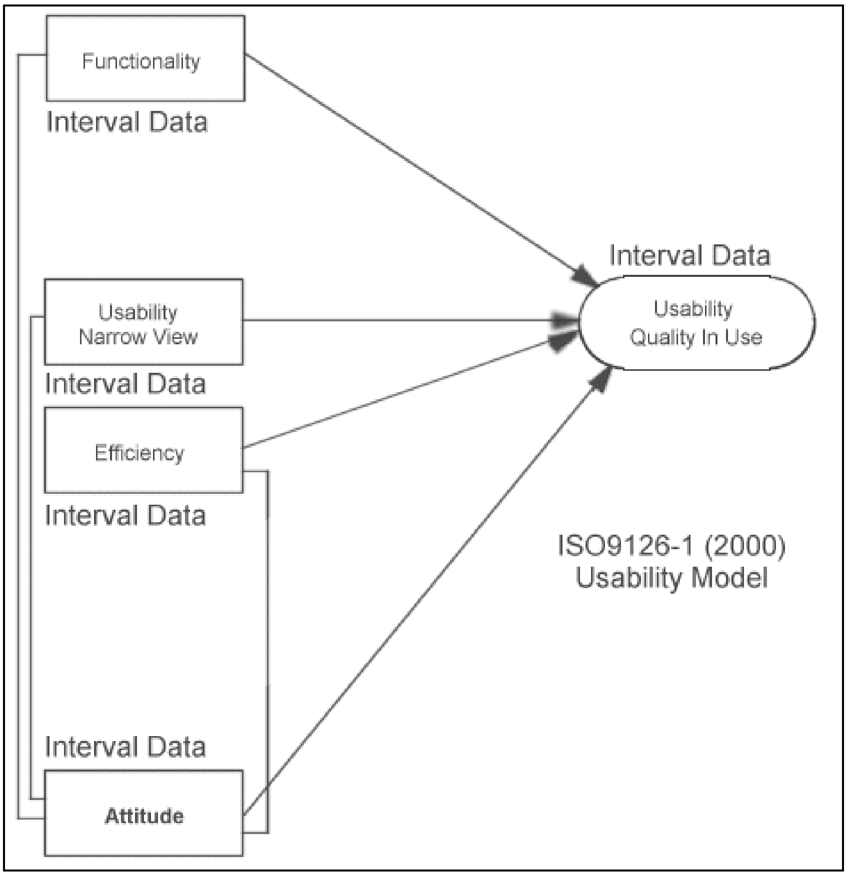


Figure 3.1: Theoretical Framework for a Context Specific Usability Study

3.2 Hypothesis of the Research

H1: Users with more positive computer attitude perceive higher system usability (quality in use).

As examined in chapter 2, computer attitude is an important concept regarding IT professionals. Any project (corporate solutions, websites, e-commerce platforms etc.) which will interact with final-end users has to take into account the existing user database with at most importance. If potential users of the system, prior to using / testing the actual system, have negative attitude towards computers, the perceived usability of the actual system can be lower. Summing up; although the system can satisfactorily achieve its design objectives, the subjective evaluation of the users, who have negative attitudes towards computers, can be negative, too. This is a very important fact that clearly acts as a bottleneck in front of the organization's interest. When the top-level management is to gather reports regarding to the tests deployed by actual system users, they have to be aware that their employees' attitudes towards computers can interfere with the reporting process. The main hypothesis of this thesis therefore focuses on this aspect; Users with more positive computer attitude perceive higher system usability.

Regarding hypothesis 2, 3 and 4 the following explanation applies. Our dependent variable usability has four independent variables of interest; functionality, time-efficiency, interface-usability and computer attitude. The first three of these independent variables are selected from ISO 9126-1 (2000) model which consists of 6 independent variables. The remaining three (reliability, maintainability and portability) are not suitable for the scenario that will be tested (shopping in an e-commerce site whose details will be supplied further in this chapter). The specific context of interest and its actual user database (e-commerce users) are not necessarily interested in the remaining three variables when evaluating an e-commerce site's usability. Reliability; a technical criteria regarding a system's reliability, maintainability; a technical criteria which is important for programmer's and system administrators, portability; a technical criteria which is again important for programmer's and system administrators.

H2: Users with more positive computer attitude perceive higher system functionality.

Hypothesis 2 examines the relationship between our independent variables; computer attitude and functionality. The designer's of the system on a software level are greatly focused on adding extra functionality to the system which will make it easier for the end-user to satisfy an objective. In our scenario of interest, our users will test drive an e-commerce site and search for information regarding cellular phones. They will compare features such as camera resolution, Bluetooth and etc. The e-commerce site which is of interest has an advanced search engine which makes it possible to directly search within features of a product. It will be interesting to see if computer attitude affects the perception of such extra functionalities as most system designers believe that better the functionality, better the perceived satisfaction levels however this may not be the case so in the existence of users with negative computer attitudes.

H3: Users with more positive computer attitude perceive higher time efficiency.

Hypothesis 3 holds on to the assumptions of hypothesis 2. In the existence of users with negative computer attitude; these users can avoid from using extra functionalities of the system which can make them complete their objectives in a shorter time period. If this is true then these users will both complete their scenarios in a longer time period and also perceived lower time efficiency. Our expectancy is that users with better computer attitude will perceive higher time-efficiency.

H4: Users with more positive computer attitude perceive better interface-usability.

Hypothesis 4 holds on to the assumptions of hypothesis 2. When we have users with positive computer attitudes they will like the interface of the system more. This is real important especially for interface designers because they always have to deal with different tastes and expectancies of different end users within the same organization. Sometimes these expectancies can be such different that it is almost impossible to find an optimum scenario which will satisfy different users in the same organization. If we are unable to falsify H4 then this has a very important practical application in the industry. Any company which is providing such design services than can first submit questionnaires to their customers which measure the existing computer attitude levels within the organization. If results point that the organization of interest is full of users with negative computer attitudes, then the service provider can choose to not serve this customer or share its findings with the top-level managers regarding to the concern that the existing user database is not suitable for objectively evaluating the interfaces of the system.

H5: Males perceive higher system usability than females.

As discussed in chapter 2, gender is a topic which has taken great interest amongst researchers. Gender and its affect on computer attitude is an area which mush research has been conducted. Hypothesis 5's objective is to examine gender and its impact upon perceived usability levels in Turkey. We expect that we will not be able to falsify H5 and males will perceive higher system usability than females. One should also note that, except the exceptions of the ex-communist countries and a single case in Hong-Kong, generally speaking, males have more positive computer attitudes than females. We believe that this will hold true for also perceived usability levels in our research.

H6: Males have more positive computer attitude than females.

Hypothesis 6 holds on to the assumptions of hypothesis 5. As explained above, the literature is full of research indicating that males have more positive computer attitude in contrast to females. We expect the same findings for Turkey. If our expectancy is found to be true, this could be due to the reason that, males are spending more time and thus have higher computer experience in contrast to females. However as has been discussed in chapter 2, it has been found that computer experience is a poor predictor of computer attitude. Summing up, this can be due to the culture of Turkey which directs males towards technically oriented tasks, which require interacting more with computers, more in contrast to females. The precise examination of this aspect is not the main objective of this research.

H7: Students who are enrolled in technical departments have more positive computer attitude than students who are enrolled in non-technical departments.

Hypothesis 7 will test to see if there are any significant differences in computer attitude between students who are studying in technical fields (such as Computer Engineering, Mechanical Engineering, Metallurgy, etc.) and non-technical fields (such as Business Administration, International Relations, etc.).

H8: Students who have successfully completed the scenario has more positive computer attitude than students who were not able to complete the scenario.

At the end of each scenario, each student fills out the research questionnaire. In the first page of the questionnaire, where the demographic information is collected, the student informs to whether they have successfully completed all objectives or not. This information is cross-checked through the database of the e-commerce system by examining that student's shopping basket which he / she have created throughout the shopping scenario. If the student has successfully added one of the five cellular phones that match the criteria of the scenario then he / she has successfully completed the scenario. Our expectation is that there will be a correlation between more positive computer attitude and successful scenario completation.

H9: The four independent variables will significantly explain the variance in overall usability.

The ninth hypothesis will make it possible to see if overall model derived from ISO 9126-1 (2000) will be able to explain significant levels of variance within overall usability.

3.3 Research Design

3.3.1 Research Objective

Primary research objective is examining the relationship between computer attitude and usability. However doing so requires careful interpretation of the literature regarding usability. As discussed in chapter 2, the broad view of usability is context dependent, meaning any situational factors impact usability such as environment, system, user, organization and etc. Whenever any context variable changes, we have a complete new scenario at hand thus we can never compare two different usability measurements directly to each other. Summing up, usability measurements can only be compared between each other if we are making the measurements on the very same context with the only changing variable being the system at hand. This is a very important aspect which has to be taken into account if one is to make direct comparison amongst different studies conducted in different contexts. In the IT industry this poses no problem as usually usability measurements are compared between different (upgraded, future) versions of the system in the same context at hand. On the contrary even within this context, the IT Company should always watch for context changes (such as employees leaving the company which were existent in the previous usability studies and new employees that has just started with the company and will participate in the current usability tests).

For successfully and accurately satisfying the primary research objective, measuring usability in a context-specific scenario is preferred whilst examining the correlations between our research variables. The specific scenario of interest is shopping in an e-commerce portal. The elements of the sample will complete a specific scenario in the selected e-commerce portal and after the study will complete the developed questionnaire. More information regarding the scenario can be found below.

3.3.2 Research Scenario: Shopping in an E-commerce Portal

The selected e-commerce portal which will be used for the field study can be examined at the internet address; www.paketticaret.com. This is an e-commerce portal which is specifically deployed for this research and thus no daily, internet traffic will be available during the field studies. This assures that the elements of the sample will not be affected from daily high traffic or server performance issues (such as long page load / processing times which can occur when the e-commerce portal attracts higher than average daily traffic and suffers from low performance). On the other hand the portal of interest is a complete functioning e-commerce portal with well over 5,000 products within the database with all product images, product specifications available. This portal is developed by "Teknolojim Software House"[*]. One can examine screen shots of the portal on the following pages. Page sequence will follow the scenario sequence so that whilst examining the screen shots of the portal, one will be able to understand the details of the scenario.

[*] Teknolojim Software Company; www.teknolojim.com

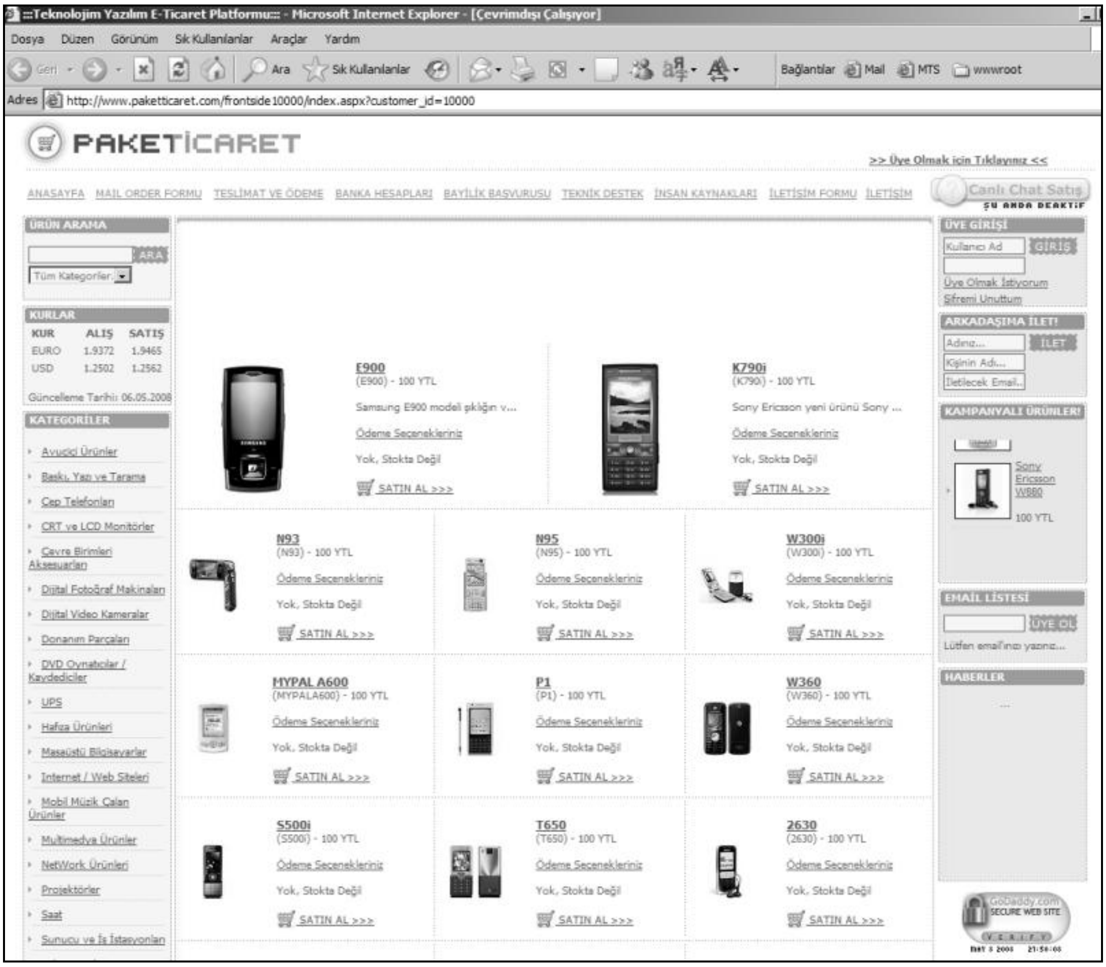


Figure 3.2: Scenario Portal’s Home Page; www.paketticaret.com

Users are welcomed by the home page in Figure 3.2. Each user is given the following same task; “I want to buy a new cellular phone. My new phone should have a camera whose resolution is 3 to 5 Mega Pixels, should have Bluetooth and should record VGA resolution (640*480 pixels) videos”. Within the cellular phones product category, the portal database contains 1034 products. Only 5 cellular phones possess these features. The users are free to choose any phone amongst these 5 phones.

On the home page, product category links are on the lower-left corner of the screen. The users should selected “Cep Telefonları” (Cellular Phones) within this part of the portal first to visit cellular phones product category page whose screen shot is available in Figure 3.3.

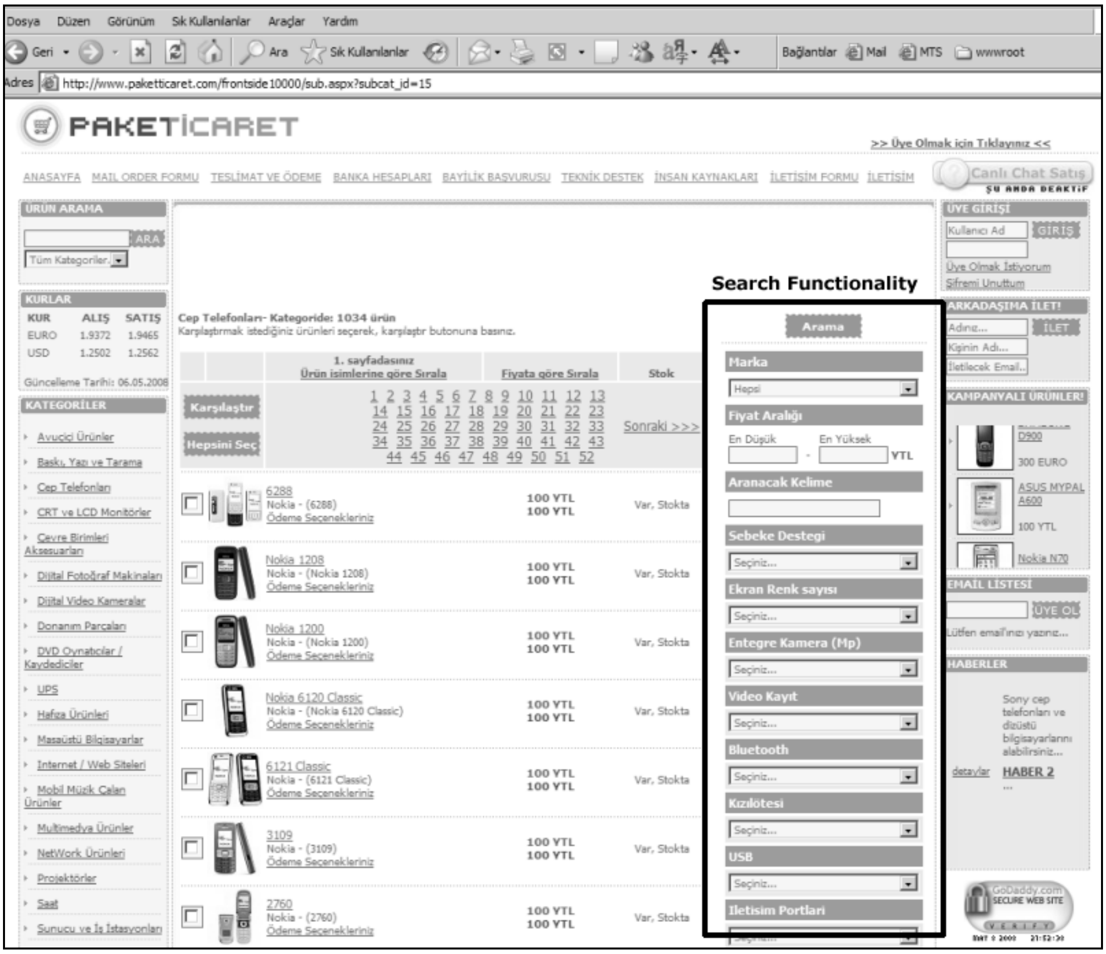
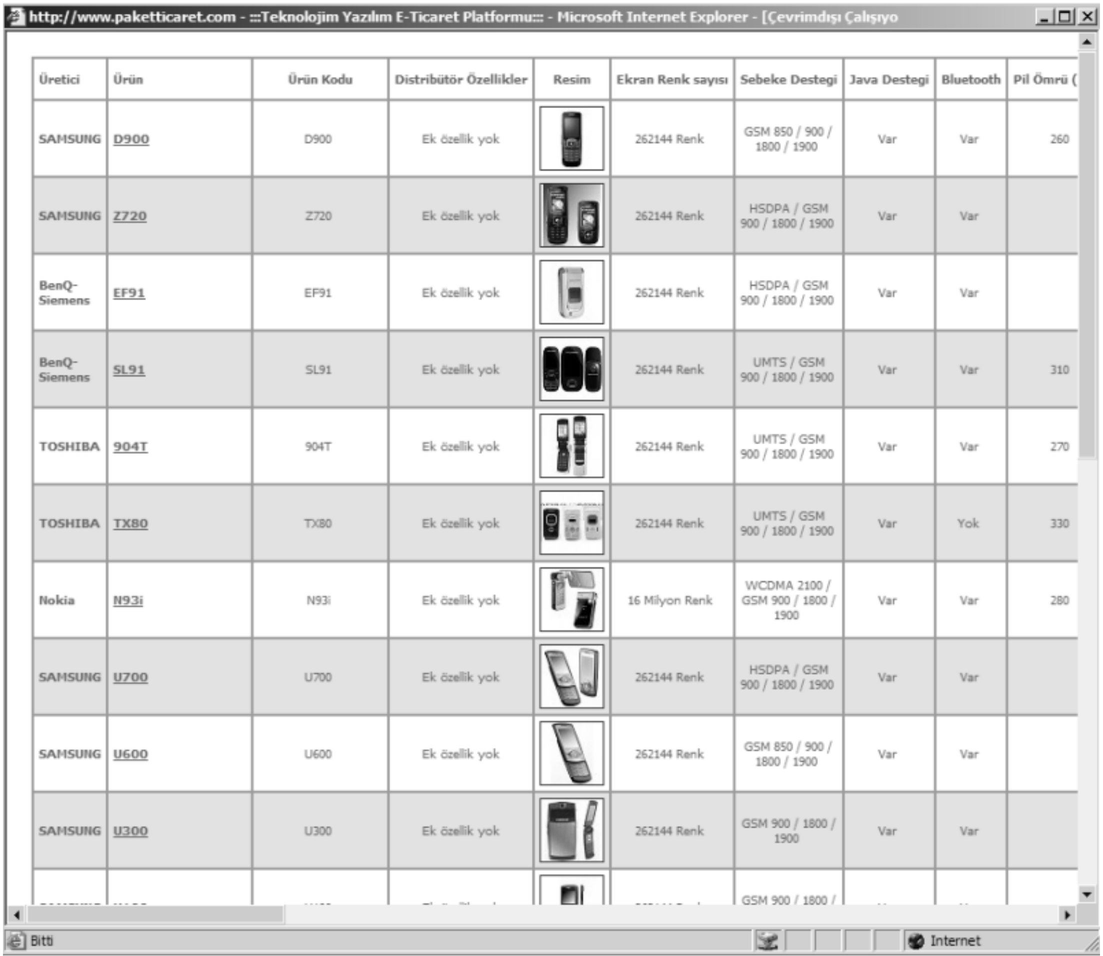


Figure 3.3: Scenario Portal’s Product Category Page for Cellular Phones; www.paketticaret.com

When the users arrive at the cellular phones product category page, they should search for information regarding to product specifications for satisfying scenario objectives. For doing this so, they can utilize built-in product search functionality of the portal. Search functionality can search through 20 different specifications of cellular phones amongst them; camera resolution, video recording resolution and Bluetooth availability are existent. Therefore users who have higher experience regarding computers and the internet will easily find search functionality and will find the 5 products which posses the specifications the scenario requires. When these 5 phones are pointed out, the users then can use “Product Comparison” feature of the portal for easily comparing these phones and their properties. The product comparison page compares each of the selected phones by their 45 different specifications and is a great tool for decision making. Its screen shot can be seen in Figure 3.4.



Üretici	Ürün	Ürün Kodu	Distribütör Özellikler	Resim	Ekran Renk sayısı	Sebeke Destegi	Java Destegi	Bluetooth	Pil Ömrü (
SAMSUNG	D900	D900	Ek özellik yok		262144 Renk	GSM 850 / 900 / 1800 / 1900	Var	Var	260
SAMSUNG	Z720	Z720	Ek özellik yok		262144 Renk	HSDPA / GSM 900 / 1800 / 1900	Var	Var	
BenQ-Siemens	EF91	EF91	Ek özellik yok		262144 Renk	HSDPA / GSM 900 / 1800 / 1900	Var	Var	
BenQ-Siemens	SL91	SL91	Ek özellik yok		262144 Renk	UMTS / GSM 900 / 1800 / 1900	Var	Var	310
TOSHIBA	904T	904T	Ek özellik yok		262144 Renk	UMTS / GSM 900 / 1800 / 1900	Var	Var	270
TOSHIBA	TX80	TX80	Ek özellik yok		262144 Renk	UMTS / GSM 900 / 1800 / 1900	Var	Yok	330
Nokia	N93i	N93i	Ek özellik yok		16 Milyon Renk	WCDMA 2100 / GSM 900 / 1800 / 1900	Var	Var	280
SAMSUNG	U700	U700	Ek özellik yok		262144 Renk	HSDPA / GSM 900 / 1800 / 1900	Var	Var	
SAMSUNG	U600	U600	Ek özellik yok		262144 Renk	GSM 850 / 900 / 1800 / 1900	Var	Var	
SAMSUNG	U300	U300	Ek özellik yok		262144 Renk	GSM 900 / 1800 / 1900	Var	Var	
[illegible]	[illegible]	[illegible]	[illegible]		[illegible]	GSM 900 / 1800 /	[illegible]	[illegible]	

Figure 3.4: Scenario Portal’s Product Comparison Page for Cellular Phones; www.paketticaret.com

In an ideal scenario, it won’t take more than 1-2 minutes for a novice user to point out 5 cellular phones of interest and use the “Product Comparison” feature to compare them amongst each other. If the user does not discover “Product Comparison” feature, which is readily available in the product category page, then he / she has to examine each phone and its specifications one at a time and will take it longer to compare all phones.

When the user finishes the comparison process, He / she will have to add his / her preferred cellular phone to shopping basket for check out. This can be done on the specific product’s page whose screen shot is in Figure 3.5.

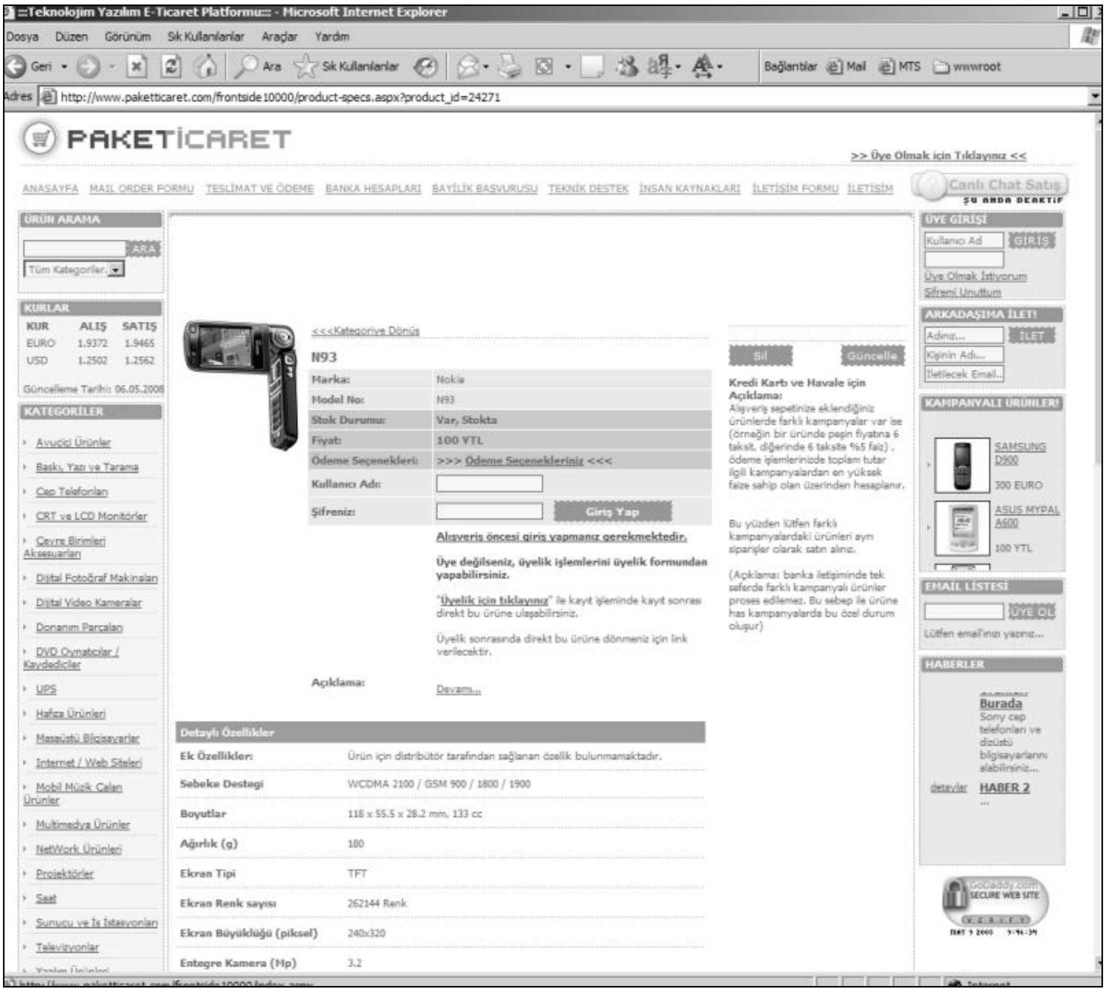


Figure 3.5: Scenario Portal’s Product Page for the Selected Phone; www.paketticaret.com

When the user has selected the specific phone which he / she would like to buy, he / she would have to add this product to the shopping basket however doing it so requires that the user has a user account within the user database of the portal. As all of the elements within the sample will be using this portal for the first time in their life times, they won’t have previously created user accounts and will have to sign up for membership before being able to add their preferred product to the shopping basket. Member registration form can be seen in Figure 3.6.

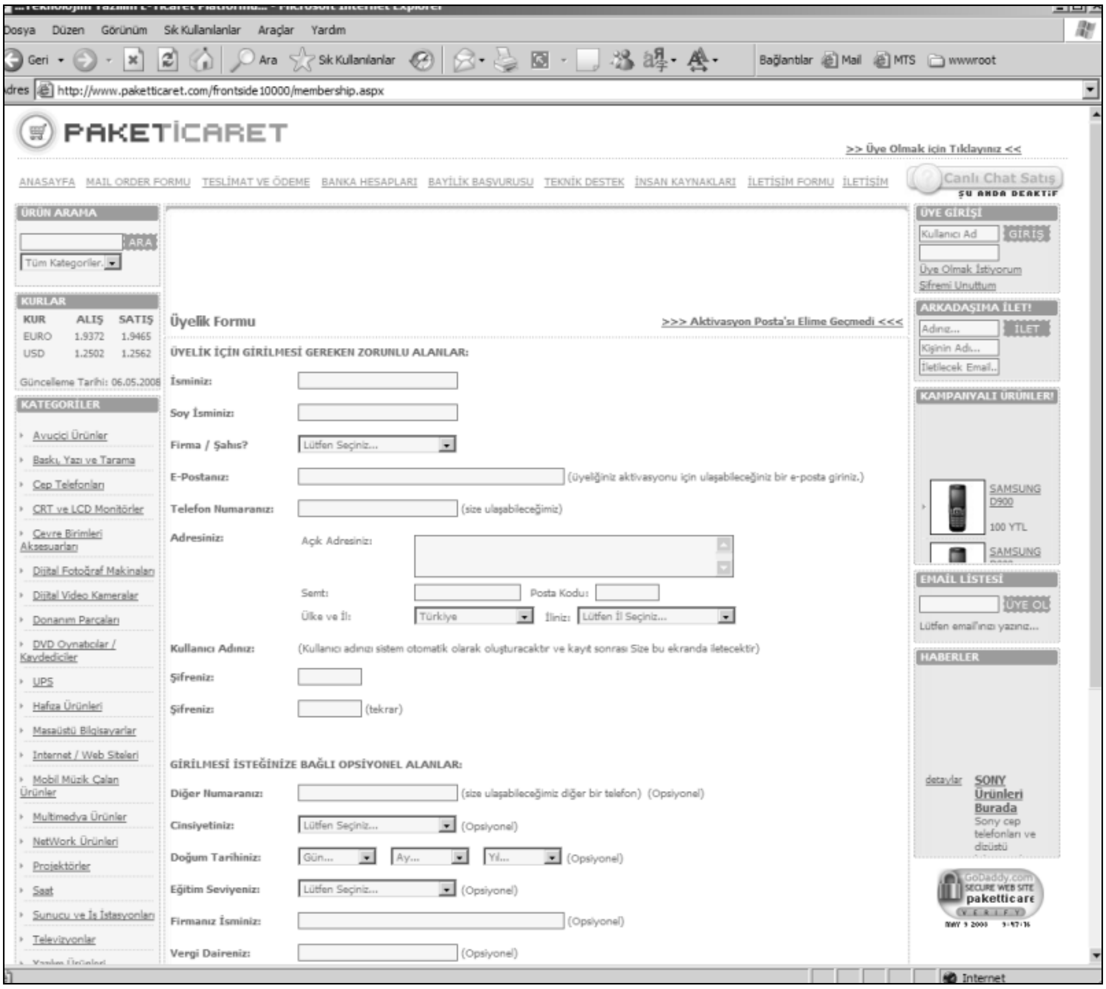


Figure 3.6: Scenario Portal's New Membership Registration Page;

www.paketticaret.com

The user will fill out each specific entry within the membership registration form for becoming a member. When they successfully fill out the form, the system will display a message that gives them their username which they will require when they are to open a session (user session). After a satisfactory membership registration process, users must have to go back to the product page where they can start a session and add their preferred product to the shopping basket. After adding their preferred product to the shopping basket, they have to go to checkout screen whose screen shot is in Figure 3.7.

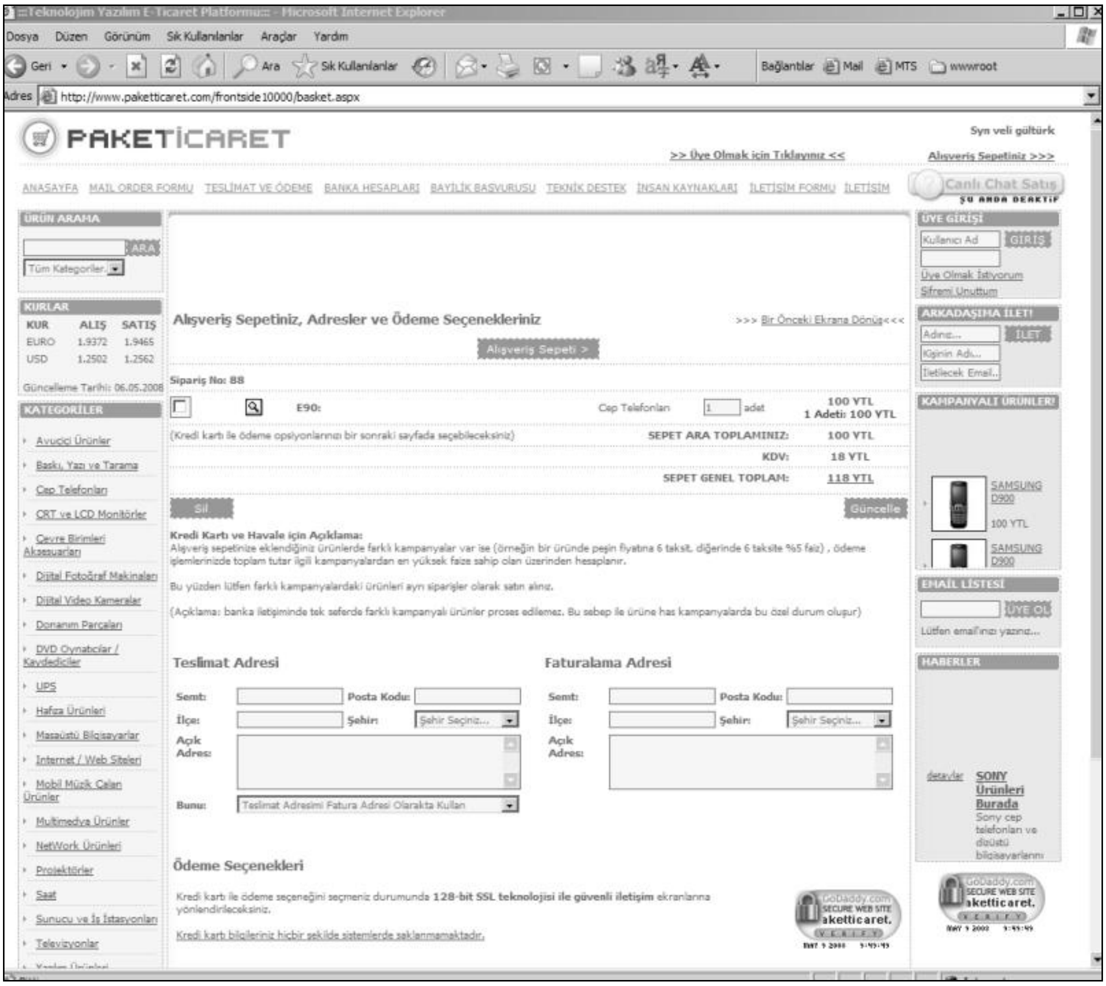


Figure 3.7: Scenario Portal's Checkout Page Which Display the Shopping Basket; www.paketticaret.com

When the users have successfully arrived at this page, with their preferred product added to their shopping basket, the scenario is over. After this, the interviewer directs the users, to the online questionnaire which they will submit.

This specific scenario which is specifically developed for this research makes it possible for researchers to control usability measurement environment (the e-commerce software, the servers, product information, etc.). Upon this scenario the specific model which is given in chapter 3.1 is developed. Within this research we are measuring the perceived usability by the internet users in the sample and thus the ISO 9126-1 (2000) model's variables are filtered to match the context. As end users do not care about technical reliability, maintainability and portability dimensions of the usability concept, these are dropped out of the model and the developed questionnaire does not measure these dimensions.

In conclusion, when a specific usability measurement is to be conducted, dimensions of the usability have to be reconsidered to match the context and unnecessary ones should be left out. One can refer back to chapter 3.1 for our theoretical framework which is a revised version of ISO 9126-1 (2000), to match the context. Computer attitude variable is this research's contribution to the model and can be applied due to the reason that, the context measures perceived usability (subjective usability evaluation of the end-user) which can be greatly affected by the user's existent attitude towards computers. However, if we were not measuring perceived usability of end-users (subjective usability) but rather objective usability (Then the interviewer's would be maybe gathering time data and other kinds of data for objective measurement) then maybe it wouldn't be that suitable to add computer attitude as a variable into the theoretical framework (Assumption is that objective performance of the users are not affected by negative computer attitude. However we do not have an idea about this phenomenon. It should be interesting to see if there are any significant objective performance differences amongst users with better computer attitude in contrast to users with negative computer attitudes).

3.3.3 Basic Research Design Issues

We will be testing 7 different hypotheses within the field study. The purpose of our research is hypothesis testing and we will be examining the correlations among computer attitude and usability (quality in use, usability from a broad perspective). The researcher interference will be minimal in the field study. Elements of the population will be completing the scenarios within their natural environments. More information regarding the field study will be supplied below. Study setting will be noncontrived. Elements of the population will complete the scenario in their daily environment (like university campus, hostel, library, internet café, etc.). This research is a field study. The unit of analysis is individuals. Every individual within the sample will test drive the scenario only once thus this is a cross-sectional study. One can examine the elements of the research design in a Table 3.1.

Elements of Research Design	
The Purpose of the Study	Hypothesis Testing
Type of Investigation	Correlational (Examining the correlations between computer attitude and usability)
Extent of Researcher Interference	Minimal
Study Setting	Noncontrived, Field Study
Unit of Analysis	Groups (Males against females… vs.)
Time Horizon	Cross-sectional

Table 3.1: Elements of Research Design

3.3.4 Sampling

The field study is a scenario which is conducted in an e-commerce portal. The population of the research is selected as the university students in Istanbul.

Istanbul is mainly selected as focus of this research due to the reason that this city is the largest city in terms of population and has the largest GDP per person amongst all cities of Turkey. Alone by itself Istanbul's population constitutes almost 20 percent of the whole population of Turkey. Almost half of the whole GDP is produced within this city. The city has the highest number of universities (26 universities, almost 204,000 students) within the country for a single city. Therefore Istanbul with its huge population and largest GDP per person is the most convenient place for a research which is on electronic commerce.

University students are primarily selected due to the reason that these students have higher computer experience in contrast to elder peoples of the society, represent both the current and future user base of the internet and are a generation which have grown with computers and internet (starting from the 1990s). These users will mostly interact with computers both in their social and professional life on a daily basis. Therefore examining their attitudes towards computers and usability have great value for supplying in-depth information to IT professionals for optimizing their solutions and services. One can quite easily assume that their computer attitudes can be more favorable in contrast to elder people who did not have the opportunity of interacting with computers to the level that they have. However this is yet to be tested on a scientific basis.

There are 26 universities educating 204,000 students in Istanbul[*]. As our population size is a couple of 100,000s, the selected sample size is 384[*2] (Sekaran, 2000). The generalizability of the research is an important criterion and as such probability sampling methods are of concern. Applying simple random sampling is not an option (as it is not feasible and the population list is not available) for a population size of 204,000. Stratified sampling is a viable option which is also more efficient in terms of examining (stratified) group relations within the population. Stratification is done on a university level. As each university's number of students is different to another, the sample also should reflect this existing fact. As a result each stratum within the population will have different number of elements within the sample. Therefore the sample will be disproportionate in between strata. The number of elements each stratum contains is calculated as; the percentage of students the university has in Istanbul X selected sample size (384). When we compute each stratum's number of

[*] Number of students, student distributions amongst universities in Istanbul are gathered from data obtained from YÖK (Yüksek Öğrenim Kurulu) (Governmental Organization who is responsible from Universities in Turkey).

[*2] Uma Sekaran suggests that for populations with sizes upto 1,000,000, samples of 384 is adequate. P. 295.

elements, some are found to be lower than 1 such as Arel and Istanbul Bilim Universities. As a matter of fact, these strata' number of elements are corrected as 1. The sample's distribution amongst universities within the population can be examined in Table 3.2 ("effective" column).

Name of University		# of Students	Ratio (%)	# of Elements	Effective
Bahçeşehir	University	5300	2,61	10	10
Beykent	University	7500	3,69	14,2	14
Doğuş	University	1550	0,76	2,9	3
Fatih	University	3500	1,72	6,6	7
Haliç	University	650	0,32	1,2	2
Işık	University	2220	1,09	4,2	4
İstanbul Bilgi	University	3800	1,87	7,2	7
İstanbul Bilim	University	370	0,18	0,7	1
İstanbul Kültür	University	5960	2,93	11,3	11
İstanbul Ticaret	University	1800	0,89	3,4	3
Kadir Has	University	1030	0,51	2	2
Koç	University	3200	1,58	6,1	6
Maltepe	University	1500	0,74	2,8	3
Okan	University	1200	0,59	2,3	2
Sabancı	University	3300	1,63	6,3	6
Yeditepe	University	2800	1,38	5,3	5
İstanbul Aydın	University	5000	2,46	9,4	9
Yakın Doğu	University	11000	5,42	20,8	21
Arel	University	0	0	0	1
Boğaziçi	University	8690	4,28	16,4	16
Galatasaray	University	2100	1,03	4	4
İstanbul Teknik	University	15100	7,44	28,6	29
İstanbul	University	51000	25,11	96,4	96
Marmara	University	41000	20,19	77,5	78
Mimar Sinan Güzel San.	University	4800	2,36	9,1	9
Yıldız Teknik	University	18700	9,21	35,4	35
TOTAL		**203070**			384

Table 3.2: Sample Distribution amongst Universities within the Population

Further sampling within each stratum is done via simple random sampling. Concluding, the sampling method deployed for this research is disproportionate stratified simple random sampling.

3.3.5 Data Collection Method

One can refer back to chapter 2 regarding different methods deployed for measuring usability. Amongst them, we are especially interested in methods that are suitable for this research's context (A method which is suitable for after scenario evaluation of the system). The research's objective is to examine the correlations between computer attitude and perceived usability (subjective usability) and therefore data will be collected from individuals (but analyzed on a group level) regarding to their perceptions. For capturing perceptions / emotions / ideas questionnaires (amongst usability inquiry methods) are already regarded as valid and dependable data collection methods.

Recalling on the theoretical framework suggested in chapter 3.2, we have to measure the following variables; functionality, interface usability, time-efficiency, computer attitude and overall usability (subjective usability). Measurement of the variables functionality, interface usability, time-efficiency and overall usability will be done via Computer System Usability Questionnaire (Lewis (1992b) CSSUQ, can be examined in chapter 2.1.4.3.6). CSSUQ is a thoroughly tested and psychometrically evaluated questionnaire which is perfect for after scenario testing. It contains 19 7-point Lickert scale items. 11 of these 19 items are found suitable for this research's context. One can find item and variable associations below (item numbers are applicable for this research's questionnaire).

- Functionality; measured by Items 3, 11, 15.
- Efficiency (Time-Efficiency); measured by Items 6, 10.
- Interface Usability; measured by Items 4, 7, 9, 12.
- Overall Usability (Broad View of Usability / Subjective Usability); measured by Items 2, 18.

Lewis found 3 factors in his original work via conducting principal factor analysis. He has named these factors as System Use (SYSUSE), Information Quality (INFOQUAL) and Interface Quality (INTERQUAL). We have examined each and every item of CSSUQ and carefully selected and then categorized the appropriate items which will measure our independent variables.

One can examine the items which are selected for measuring functionality below.

- I can effectively complete my work using this system. (Item 3)
- The information is effective in helping me complete my work. (Item 11)
- This system has all the functions and capabilities I expect it to have. (Item 15)

One can examine the items which are selected for measuring efficiency (time-efficiency) below.

- I am able to complete my work quickly using this system. (Item 6)
- I am able to efficiently complete my work using this system. (Item 10)

One can examine the items which are selected for measuring interface-usability below.

- It is simple to use this system. (Item 4)
- The system gives error messages that clearly tell me how to fix problems. (Item 7)
- It is easy to find the information I need. (Item 9)
- The information (such as on-line help, on-screen messages and other documentation) provided with this system is clear. (Item 12)

We believe that interface quality has priority over functionality for a user who is to evaluate a specific site. However one should not also forget that extra value-adding functionality (such as advanced search engines, compare product options and etc.) is also welcomed by customers and improve their perceptions of the site. As a result these users can find the specific site easier-to-use due to these extra functionalities. However if interface quality is not existent, having extra functionalities does not count. Therefore interface quality is a prerequisite to a site which is easy-to-use.

The word “this system”, existent in items in original CSSUQ, is changed to “this e-commerce site” within the final developed questionnaire as it is more suitable for the context.

The independent variable, computer attitude, is measured by items 1, 5, 8, 14 and 17. These items are taken from Loyd and Gressard (1984)’s Computer Attitude Measurement scale which has been examined in chapter 2.2.3.1. This scale contains Lickert-scale items.

- I like working with computers. (Item 1)
- I look forward to those aspects of my job that require me to use a computer. (Item 5)
- I get bored quickly when working on a computer. (Item 8)
- Once I start working on the computer, I find it hard to stop. (Item 14)
- Using a computer is frustrating for me. (Item 17)

The dependent variable, overall usability, is measured by items 2 and 18. This scale contains Lickert-scale items.

- Overall, I am satisfied with the simplicity of this e-commerce site. (Item 2)
- Overall, I am satisfied with this e-commerce site. (Item 18)

Original wording of the items in this scale is not changed to reflect the scenario as we intend to measure the user's general attitude towards computers, not specifically towards e-commerce sites.

Items 8 and Item 17 will be reverse scored due to negative wording when data is to be inserted into SPSS.

All independent variables and the dependent variables will be measured on an interval scale therefore making it possible to conduct parametric statistical tests (for testing hypothesis 1, 2, 3 and 4). Male and female information will be on nominal scale (hypothesis 5 and 6 are related with gender). Within the online questionnaire's first part, where we are collecting demographic data, the participants will supply information in regards to their department within their universities. After completion of the field study, when we are to import this data to SPSS, this department information will be transformed into a nominal scale. This process will be done with the classification of each department in regards to whether they are technical disciplines (such as Computer engineering) or non-technical disciplines (such as International Relations). This will make it possible to test hypothesis 7.

Developed questionnaire can be examined in Appendix A.

3.4 Research Results

Results of the research can be found within the following sections.

3.4.1 Internal Consistency

The interitem consistency reliability has been computed for four of our independent variables and our dependent variable. As can be seen all are above 0.78.

Variable	# of Items	Alpha
Functionality	3	.78
Efficiency	2	.876
Interface Usability	4	.904
Computer Attitude	5	.821
Overall Usability	2	.819

Table 3.3: Computed Cronbach's Alpha for Research Variables

The internal consistency reliability of the measures used in this study can be considered to be good.

3.4.2 Descriptive Statistics

One can find the frequency distributions for important demographic variables of the research below. The measures of central tendencies are also supplied in the following subsection.

3.4.2.1 Frequency Distributions

From the frequencies obtained for our research variables, the following are found. 54.4 percent of the respondents are males ("1") and 45.6 percent are females ("0").

Statistics

Gender

N	Valid	384
	Missing	0

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	175	45,6	45,6	45,6
	1,00	209	54,4	54,4	100,0
	Total	384	100,0	100,0	

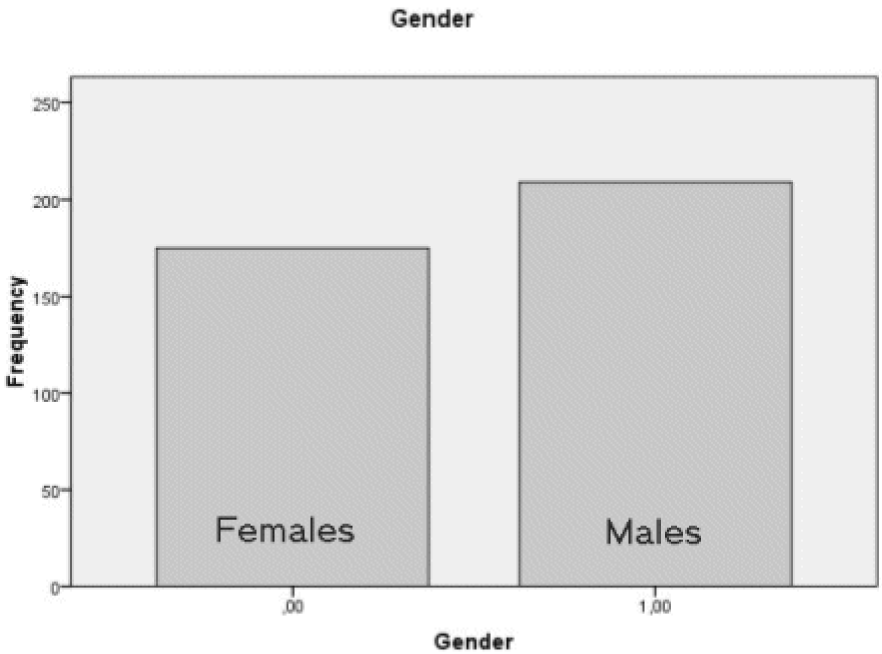


Figure 3.8: Frequency Distribution for Gender

44.5 percent of the respondents are studying in technical departments whereas 55.5 percent of the respondents are studying in non-technical departments.

Statistics

TechnicalvsNonTechnical

N	Valid	384
	Missing	0

TechnicalvsNonTechnical

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	213	55,5	55,5	55,5
	1,00	171	44,5	44,5	100,0
	Total	384	100,0	100,0	

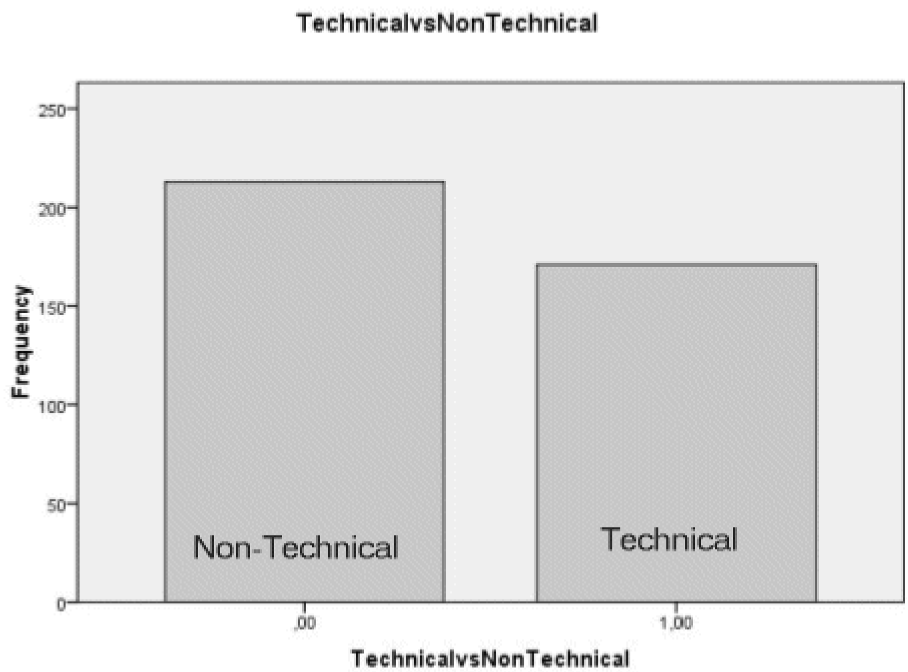


Figure 3.9: Frequency Distribution for Departments

In Figure 3.9; “1” represents “Technical Departments” category and “0” represents “Non Technical Departments”.

Functionality

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	1	,3	,3	,3
	1,67	2	,5	,5	,8
	2,00	6	1,6	1,6	2,3
	2,33	5	1,3	1,3	3,6
	2,67	2	,5	,5	4,2
	3,33	4	1,0	1,0	5,2
	3,67	8	2,1	2,1	7,3
	4,00	18	4,7	4,7	12,0
	4,33	18	4,7	4,7	16,7
	4,67	28	7,3	7,3	24,0
	5,00	41	10,7	10,7	34,6
	5,33	58	15,1	15,1	49,7
	5,67	58	15,1	15,1	64,8
	6,00	47	12,2	12,2	77,1
	6,33	53	13,8	13,8	90,9
	6,67	23	6,0	6,0	96,9
	7,00	12	3,1	3,1	100,0
	Total	384	100,0	100,0	

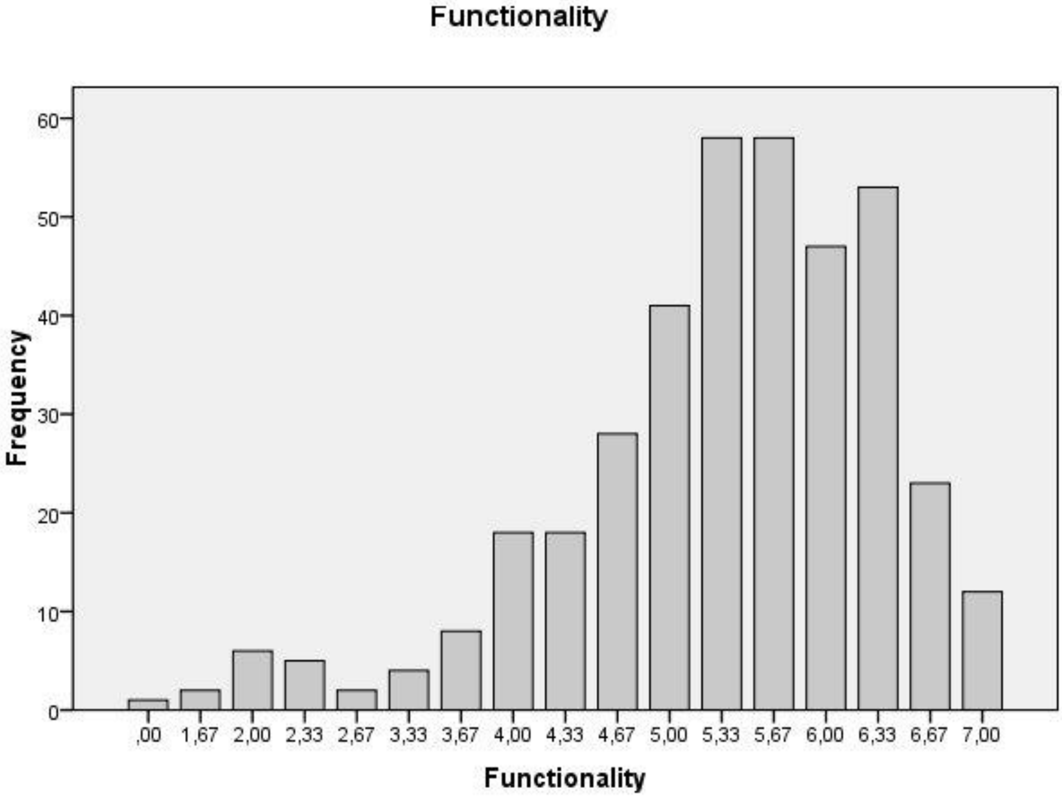


Figure 3.10: Frequency Distribution for Functionality

The frequency distribution for functionality seems to be normal (Figure 3.10). Statistical tests for testing the goodness of fit will be supplied later.

Efficiency

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	1	,3	,3	,3
	1,00	2	,5	,5	,8
	1,50	5	1,3	1,3	2,1
	2,00	8	2,1	2,1	4,2
	3,00	6	1,6	1,6	5,7
	3,50	9	2,3	2,3	8,1
	4,00	12	3,1	3,1	11,2
	4,50	35	9,1	9,1	20,3
	5,00	48	12,5	12,5	32,8
	5,50	30	7,8	7,8	40,6
	6,00	99	25,8	25,8	66,4
	6,50	46	12,0	12,0	78,4
	7,00	83	21,6	21,6	100,0
	Total	384	100,0	100,0	

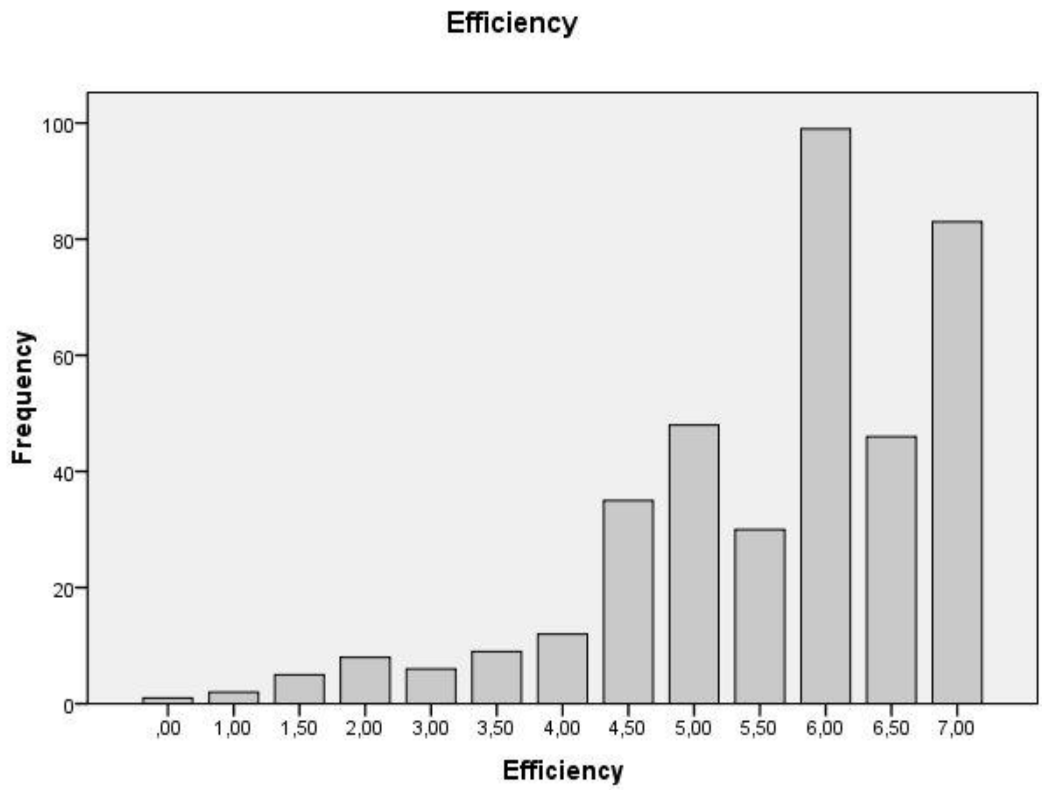


Figure 3.11: Frequency Distribution for Efficiency

The frequency distribution for efficiency seems to be negatively skewed (Figure 3.11).

InterfaceUsability

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	1	,3	,3	,3
	1,00	2	,5	,5	,8
	1,75	3	,8	,8	1,6
	2,00	1	,3	,3	1,8
	2,25	4	1,0	1,0	2,9
	2,50	4	1,0	1,0	3,9
	2,75	2	,5	,5	4,4
	3,00	6	1,6	1,6	6,0
	3,25	4	1,0	1,0	7,0
	3,50	11	2,9	2,9	9,9
	3,75	2	,5	,5	10,4
	4,00	6	1,6	1,6	12,0
	4,25	11	2,9	2,9	14,8
	4,50	11	2,9	2,9	17,7
	4,75	24	6,2	6,2	24,0
	5,00	25	6,5	6,5	30,5
	5,25	53	13,8	13,8	44,3
	5,50	32	8,3	8,3	52,6
	5,75	44	11,5	11,5	64,1
	6,00	47	12,2	12,2	76,3
	6,25	39	10,2	10,2	86,5
	6,50	19	4,9	4,9	91,4
	6,75	17	4,4	4,4	95,8
	7,00	16	4,2	4,2	100,0
	Total	384	100,0	100,0	

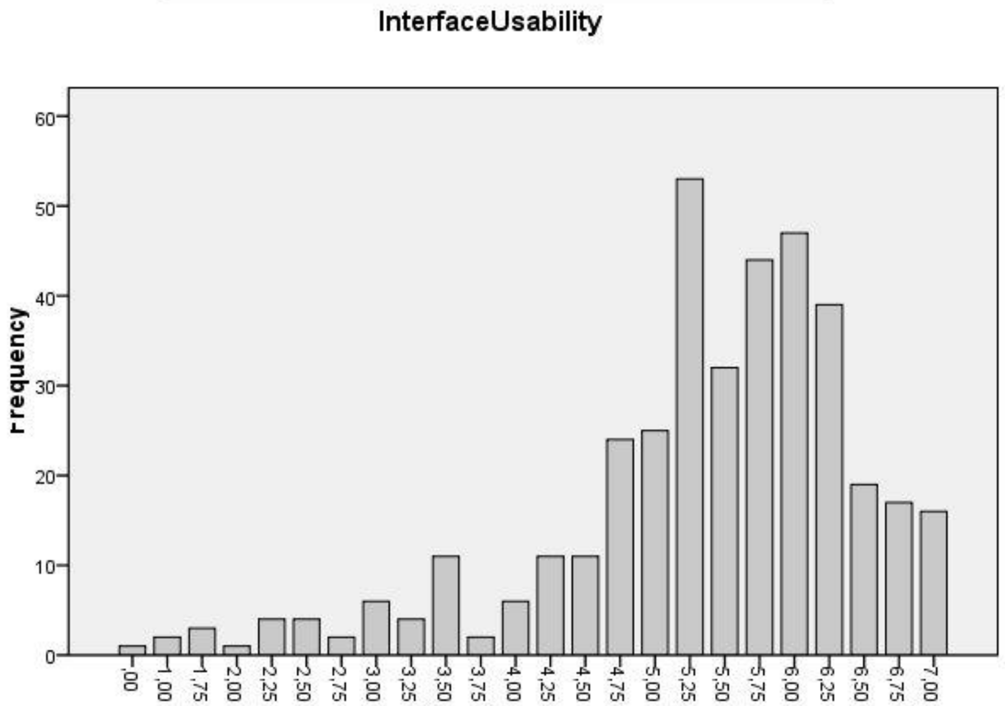


Figure 3.12: Frequency Distribution for Interface Usability

The frequency distribution for interface-usability seems to be slightly negatively skewed (Figure 3.12).

Attitude

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	1	,3	,3	,3
	1,00	2	,6	,6	,9
	2,60	2	,6	,6	1,4
	2,80	2	,6	,6	2,0
	3,00	5	1,4	1,4	3,5
	3,20	3	,9	,9	4,3
	3,40	6	1,7	1,7	6,1
	3,60	7	2,0	2,0	8,1
	3,80	6	1,7	1,7	9,8
	4,00	23	6,6	6,6	16,4
	4,20	17	4,9	4,9	21,3
	4,40	12	3,5	3,5	24,8
	4,60	16	4,6	4,6	29,4
	4,80	20	5,8	5,8	35,2
	5,00	13	3,7	3,7	38,9
	5,20	11	3,2	3,2	42,1
	5,40	24	6,9	6,9	49,0
	5,60	19	5,5	5,5	54,5
	5,80	21	6,1	6,1	60,5
	6,00	21	6,1	6,1	66,6
	6,20	37	10,7	10,7	77,2
	6,40	21	6,1	6,1	83,3
	6,60	21	6,1	6,1	89,3
	6,80	19	5,5	5,5	94,8
	7,00	18	5,2	5,2	100,0
	Total	347	100,0	100,0	

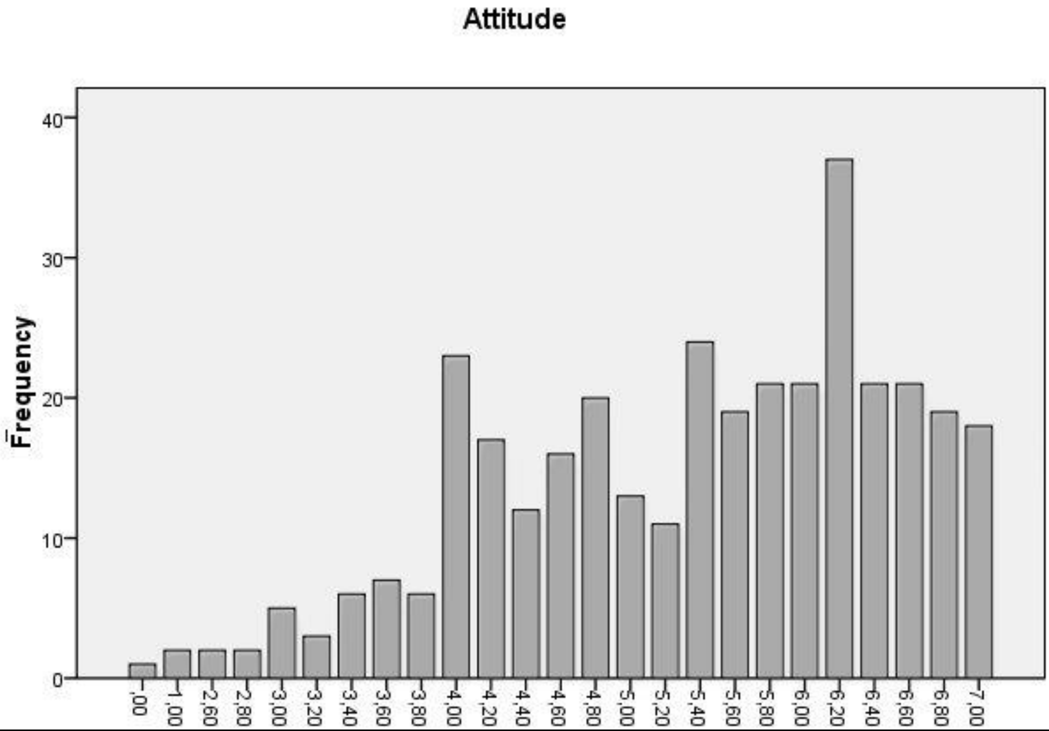


Figure 3.13: Frequency Distribution for Computer Attitude

The frequency distribution for computer attitude seems to be slightly negatively skewed (Figure 3.13).

OverallUsability

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	1	,3	,3	,3
	1,00	2	,5	,5	,8
	2,00	6	1,6	1,6	2,3
	2,50	6	1,6	1,6	3,9
	3,00	12	3,1	3,1	7,0
	3,50	10	2,6	2,6	9,6
	4,00	27	7,0	7,0	16,7
	4,50	30	7,8	7,8	24,5
	5,00	67	17,4	17,4	41,9
	5,50	71	18,5	18,5	60,4
	6,00	80	20,8	20,8	81,2
	6,50	52	13,5	13,5	94,8
	7,00	20	5,2	5,2	100,0
	Total	384	100,0	100,0	

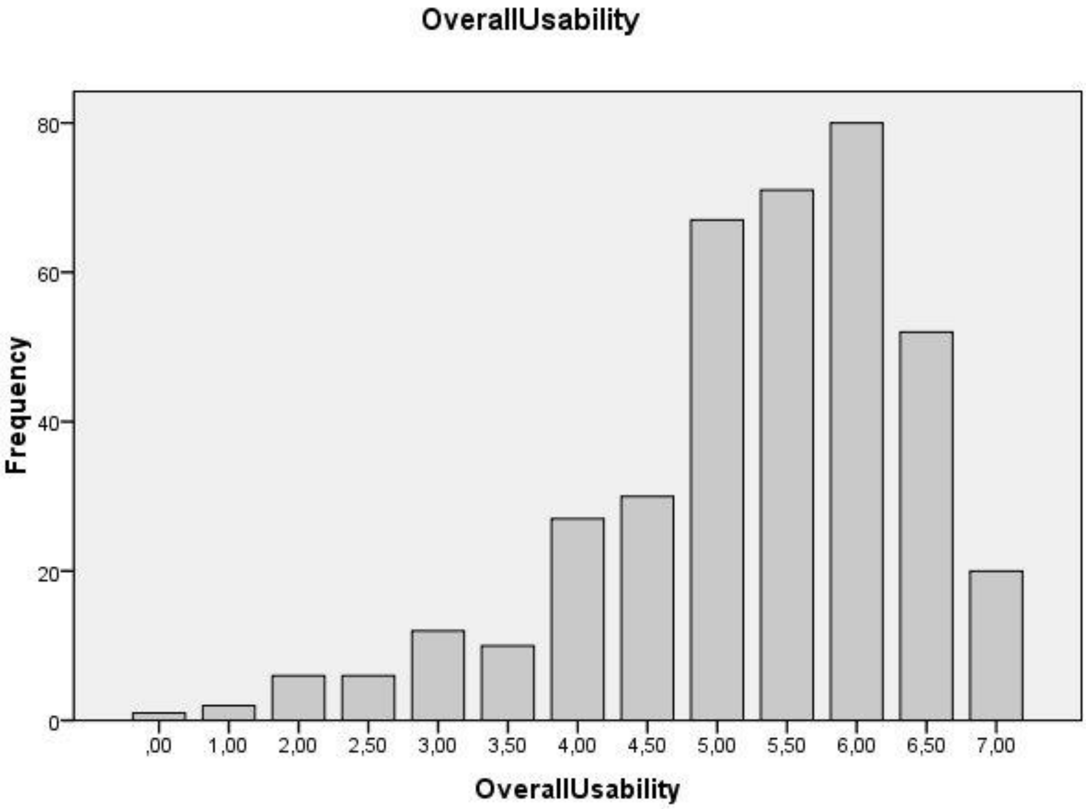


Figure 3.14: Frequency Distribution for Usability (overall usability)

The frequency distribution for overall usability seems to be normal (Figure 3.14).

3.4.2.2 Measures of Central Tendencies and Dispersion

The maximum, minimum, means, standard deviations and variance are obtained for all of our variables (Table 3.4).

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Functionality	384	,00	7,00	5,3481	1,07295	1,151
Efficiency	384	,00	7,00	5,6237	1,30061	1,692
Attitude	384	,00	7,00	5,3443	1,18508	1,404
InterfaceUsability	384	,00	7,00	5,3470	1,14280	1,306
OverallUsability	384	,00	7,00	5,2773	1,16376	1,354
Valid N (listwise)	384					

Table 3.4: Research Variables Descriptive Statistics

Both independent variables and dependent variable were tapped on a seven-point Lickert scale and thus all are interval. The results show that the functionality of the e-commerce system in the scenario was high (5.32). The efficiency (time-efficiency) of the system perceived by respondents was real high (5.62). The interface usability of the system was also high (5.33). The overall usability of the system was high (5.28). The computer attitudes of the respondents were high (5.35). The maximum of 7 indicates that computer attitude is very favorable whereas the minimum of 1 indicates that the sample has real negative feelings towards computers.

The variance for functionality, efficiency, interface usability and overall usability was low (around 1.1). This indicates that all respondents are close to the mean on these variables.

Concluding, the sample found the e-commerce system's usability (overall usability) high. Computer attitude within the sample was high.

3.4.2.3 Kolmogorov-Smirnoff Tests

Testing for normality is important for the research for being able to use the parametric-tests for analyzing our data. The one-sample Kolmogorov-Smirnoff test tests if the distribution of the sample is different or not from a normal distribution. Results can be found below for all of the research variables.

One-Sample Kolmogorov-Smirnov Test

		Functionality
N		384
Normal Parameters[a]	Mean	5,3481
	Std. Deviation	1,07295
Most Extreme Differences	Absolute	,148
	Positive	,088
	Negative	-,148
Kolmogorov-Smirnov Z		2,903
Asymp. Sig. (2-tailed)		,000

a. Test distribution is Normal.

Figure 3.15: Goodness of Fit Test for Functionality

The result indicates that the sample distribution for functionality is not different from a normal distribution (Figure 3.15).

One-Sample Kolmogorov-Smirnov Test

		Efficiency
N		384
Normal Parameters[a]	Mean	5,6237
	Std. Deviation	1,30061
Most Extreme Differences	Absolute	,208
	Positive	,145
	Negative	-,208
Kolmogorov-Smirnov Z		4,068
Asymp. Sig. (2-tailed)		,000

a. Test distribution is Normal.

Figure 3.16: Goodness of Fit Test for Efficiency

The result indicates that the sample distribution for efficiency is not different from a normal distribution (Figure 3.16).

One-Sample Kolmogorov-Smirnov Test

		Interface Usability
N		384
Normal Parameters[a]	Mean	5,3470
	Std. Deviation	1,14280
Most Extreme Differences	Absolute	,161
	Positive	,079
	Negative	-,161
Kolmogorov-Smirnov Z		3,165
Asymp. Sig. (2-tailed)		,000

a. Test distribution is Normal.

Figure 3.17: Goodness of Fit Test for Interface Usability

The result indicates that the sample distribution for interface usability is not different from a normal distribution (Figure 3.17).

One-Sample Kolmogorov-Smirnov Test

		Attitude
N		384
Normal Parameters[a]	Mean	5,3443
	Std. Deviation	1,18508
Most Extreme Differences	Absolute	,108
	Positive	,081
	Negative	-,108
Kolmogorov-Smirnov Z		2,117
Asymp. Sig. (2-tailed)		,000

a. Test distribution is Normal.

Figure 3.18: Goodness of Fit Test for Computer Attitude

The result indicates that the sample distribution for computer attitude is not different from a normal distribution (Figure 3.18).

One-Sample Kolmogorov-Smirnov Test

		Overall Usability
N		384
Normal Parameters[a]	Mean	5,2773
	Std. Deviation	1,16376
Most Extreme Differences	Absolute	,161
	Positive	,095
	Negative	-,161
Kolmogorov-Smirnov Z		3,155
Asymp. Sig. (2-tailed)		,000

a. Test distribution is Normal.

Figure 3.19: Goodness of Fit Test for Usability (overall usability)

The result indicates that the sample distribution for overall usability is not different from a normal distribution (Figure 3.19).

Concluding, sample distribution can be considered to be normal and therefore the usage of parametric tests are justified.

3.4.3 Inferential Statistics

Pearson correlation matrix and hypothesis testing subsections can be examined below.

3.4.3.1 Pearson Correlation

The Pearson correlation matrix obtained for our five interval-scaled variables is shown in Table 3.5.

		Functionality	Efficiency	Attitude	Interface Usability	Overall Usability
Functionality	Pearson Correlation	1,000	,783**	,566**	,853**	,840**
	Sig. (2-tailed)		,000	,000	,000	,000
	N	384,000	384	384	384	384
Efficiency	Pearson Correlation	,783**	1,000	,582**	,803**	,725**
	Sig. (2-tailed)	,000		,000	,000	,000
	N	384	384,000	384	384	384
Attitude	Pearson Correlation	,566**	,582**	1,000	,549**	,487**
	Sig. (2-tailed)	,000	,000		,000	,000
	N	384	384	384,000	384	384
InterfaceUsability	Pearson Correlation	,853**	,803**	,549**	1,000	,832**
	Sig. (2-tailed)	,000	,000	,000		,000
	N	384	384	384	384,000	384
OverallUsability	Pearson Correlation	,840**	,725**	,487**	,832**	1,000
	Sig. (2-tailed)	,000	,000	,000	,000	
	N	384	384	384	384	384,000

**. Correlation is significant at the 0.01 level (2-tailed).

Table 3.5: Pearson Correlation Matrix for Research Variables

All correlations are significant at the level 0.01. The results indicate that overall usability, as expected, is positively correlated with all of our independent variables. That is overall usability perception is high if functionality is high, time-efficiency is high and interfaces of the software are user-friendly. Computer attitude is also moderately positively correlated (.549) with perceived usability. As expected; users with more favorable/positive computer attitude perceive higher system usability levels.

Functionality, Efficiency and Interface Usability are highly correlated (around .8) with overall usability and thus the validity of these items have to be questioned. When examined each of these variables and their associated items within the questionnaire, it is concluded that each of these items are measuring different aspects of the usability concept. However within the research scenario of shopping in an e-commerce portal, sample elements would have perceived functionality highly correlated with interface usability. Thus it would have been possible that the elements evaluated functionality as the same thing as interface usability. The very high correlation amongst these two variables (.853, significant at the level .01) is an indication for this very fact. Maybe functionality aspect of the concept is more appropriate for evaluations of corporate software where users are to make decisions on the very functionality of the product when they are to decide. They are aware of the very distinct categorization of a user-friendly product (which is easy to operate) and a functional product (which contains the specific functionalities that are expected). A product can be functional however not user-friendly. In the context of internet environments, it is possible that this distinct categorization of these two variables can be somewhat blurred.

This evaluation can also be valid for the high correlation amongst efficiency and interface usability (.803). In the corporate buying patterns, it is possible that users can find a specific software fast to operate but they would not like the visual properties of the interface. For internet users this distinct evaluation pattern can also be quite blurred, again. However this assumption is also an untested fact which is yet to be researched.

3.4.3.2 Hypothesis Testing

We have generated nine hypotheses for this study. Hypothesis 1, 2, 3, 4, 5, 6, 7 and 8 call for a t-test. Hypothesis 9 calls for a multiple regression analysis.

3.4.3.2.1 Hypothesis 1

The null and alternate statements for hypothesis 1 are as follows:

H1o: There will be no difference between users with more favorable/positive computer attitudes and less favorable/negative attitudes in their perceived usability (quality in use).

H1a: Users with more positive computer attitude perceive higher system usability (quality in use).

Group Statistics

	Grou pin...	N	Mean	Std. Deviation	Std. Error Mean
OverallUsability	1	223	5,6457	,97725	,06544
	0	161	4,7671	1,21067	,09541

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
OverallUsability	Equal variances assumed	7,429	,007	7,858	382	,000	,87866	,11181	,65881	1,09851
	Equal variances not assumed			7,594	298,355	,000	,87866	,11570	,65097	1,10635

Table 3.6: T-Test Output for Hypothesis 1

Table 3.6 demonstrates the difference in means of 5.65 and 4.77 with standard deviations of 0.98 and 1.21 is significant at the level of 0.007. Therefore null hypothesis can be rejected. That is there are significant differences between users with more favorable/positive computer attitude and less favorable/negative computer attitude in their perceived usability.

3.4.3.2.2 Hypothesis 2

The null and alternate statements for hypothesis 2 are as follows:

H2o: There will be no difference between users with more favorable/positive computer attitudes and less favorable/negative attitudes in their perceived functionality.

H2a: Users with more positive computer attitude perceive higher system functionality.

Table 3.7 presents the difference in means of 5.73 and 4.82 with standard deviations of 0.81 and 1.17 is significant at the level of 0.000 (rounded). Therefore null hypothesis can be rejected. That is there are significant differences between users with more favorable/positive computer attitude and less favorable/negative computer attitude in their perceived functionality.

Group Statistics

	Grou pin...	N	Mean	Std. Deviation	Std. Error Mean
Functionality	1	223	5,7250	,80978	,05423
	0	161	4,8261	1,17206	,09237

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Functionality	Equal variances assumed	15,314	,000	8,887	382	,000	,89888	,10114	,70001	1,09774
	Equal variances not assumed			8,392	266,475	,000	,89888	,10711	,68798	1,10977

Table 3.7: T-Test Output for Hypothesis 2

3.4.3.2.3 Hypothesis 3

The null and alternate statements for hypothesis 3 are as follows:

H3o: There will be no difference between users with more favorable/positive computer attitudes and less favorable/negative attitudes in their perceived efficiency (time-efficiency).

H3a: Users with more positive computer attitude perceive higher efficiency (time-efficiency).

As can be seen in Table 3.8 the difference in means of 6.12 and 4.93 with standard deviations of 0.93 and 1.42 is significant at the level of 0.000 (rounded). Therefore null hypothesis can be rejected. That is there are significant differences between users with more favorable/positive computer attitude and less favorable/negative computer attitude in their perceived efficiency (time-efficiency).

Group Statistics

	Grou pin...	N	Mean	Std. Deviation	Std. Error Mean
Efficiency	1	223	6,1233	,93327	,06250
	0	161	4,9317	1,41807	,11176

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Efficiency	Equal variances assumed	21,070	,000	9,923	382	,000	1,19164	,12009	,95552	1,42777
	Equal variances not assumed			9,306	257,560	,000	1,19164	,12805	,93949	1,44379

Table 3.8: T-Test Output for Hypothesis 3

3.4.3.2.4 Hypothesis 4

The null and alternate statements for hypothesis 4 are as follows:

H4o: There will be no difference between users with more favorable/positive computer attitudes and less favorable/negative attitudes in their perceived interface-usability.

H4a: Users with more positive computer attitude perceive higher interface-usability.

As can be seen in Table 3.9 the difference in means of 5.74 and 4.81 with standard deviations of 0.86 and 1.26 is significant at the 0.000 level (rounded). Therefore null hypothesis can be rejected. That is there are significant differences between users with more favorable/positive computer attitude and less favorable/negative computer attitude in their perceived interface usability.

Group Statistics
Group Statistics

	Grouping...	N	Mean	Std. Deviation	Std. Error Mean
InterfaceUsability	1	223	5,7354	,86216	,05773
	0	161	4,8090	1,26229	,09948

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
InterfaceUsability	Equal variances assumed	18,642	,000	8,544	382	,000	,92642	,10844	,71322	1,13962
	Equal variances not assumed			8,054	264,317	,000	,92642	,11502	,69994	1,15289

Table 3.9: T-Test Output for Hypothesis 4

3.4.3.2.5 Hypothesis 5

The null and alternate statements for hypothesis 5 are as follows:

H5o: There will be no difference between males and females in their perceived system usability (overall usability).

H5a: Males perceive higher system usability than females.

As can be seen in Table 3.10 the difference in means of 5.18 and 5.39 with standard deviations of 1.15 and 1.7 is not significant. Therefore null hypothesis can not be rejected. That is there are not significant differences between males and females in their perceived system usability (overall usability). This is contrary to our expectations.

Group Statistics

	Gender	N	Mean	Std. Deviation	Std. Error Mean
OverallUsability	1,00	209	5,1794	1,15092	,07961
	,00	175	5,3943	1,17149	,08856

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
OverallUsability	Equal variances assumed	,873	,351	-1,807	382	,072	-,21486	,11889	-,44863	,01891
	Equal variances not assumed			-1,804	367,887	,072	-,21486	,11908	-,44902	,01930

Table 3.10: T-Test Output for Hypothesis 5

3.4.3.2.6 Hypothesis 6

The null and alternate statements for hypothesis 6 are as follows:

H6o: There will be no difference between males and females in their computer attitude.

H6a: Males have more positive computer attitude than females.

As can be seen in Table 3.11 the difference in means of 5.38 and 5.3 with standard deviations of 1.23 and 1.13 is not significant. Therefore null hypothesis can not be rejected. That is there are not significant differences between males and females in their computer attitude. This is contrary to our expectations.

Group Statistics

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Attitude	1,00	209	5,3818	1,22876	,08500
	,00	175	5,2994	1,13256	,08561

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Attitude	Equal variances assumed	1,136	,287	,678	382	,498	,08239	,12151	-,15653	,32131
	Equal variances not assumed			,683	378,466	,495	,08239	,12064	-,15482	,31960

Table 3.11: T-Test Output for Hypothesis 6

3.4.3.2.7 Hypothesis 7

The null and alternate statements for hypothesis 7 are as follows:

H7o: There will be no difference between students who are studying in technical departments and who are studying in non-technical departments in their computer attitude.

H7a: Students who are enrolled in technical departments have more positive computer attitude than students who are enrolled in non-technical departments.

As can be seen in Table 3.12 the difference in means of 5.61 and 5.13 with standard deviations of 1.16 and 1.16 is not significant. Therefore null hypothesis can be rejected. That is there are significant differences between students who are studying in technical departments and who are studying in non-technical departments in their computer attitude.

Group Statistics

	Technic...	N	Mean	Std. Deviation	Std. Error Mean
Attitude	1,00	171	5,6164	1,15757	,08852
	,00	213	5,1258	1,16406	,07976

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Attitude	Equal variances assumed	1,098	,295	4,114	382	,000	,49055	,11923	,25613	,72498
	Equal variances not assumed			4,117	365,108	,000	,49055	,11915	,25624	,72487

Table 3.12: T-Test Output for Hypothesis 7

3.4.3.2.8 Hypothesis 8

The null and alternate statements for hypothesis 8 are as follows:

H8o: There will be no difference between students who have successfully completed the scenario and who were not able to complete the scenario in their computer attitude.

H8a: Students who have successfully completed the scenario has more positive computer attitude than students who were not able to complete the scenario.

As can be seen in Table 3.13 the difference in means of 5.44 and 3.82 with standard deviations of 1.1 and 1.2 is not significant. Therefore null hypothesis can be rejected. That is there are significant differences between students who have successfully completed the scenario and who were not able to complete the scenario in their computer attitude.

Group Statistics

	Scenari...	N	Mean	Std. Deviation	Std. Error Mean
Attitude	1,00	361	5,4410	1,11599	,05874
	,00	23	3,8261	1,22742	,25593

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Attitude	Equal variances assumed	,000	,982	6,689	382	,000	1,61491	,24144	1,14019	2,08963
	Equal variances not assumed			6,150	24,374	,000	1,61491	,26259	1,07339	2,15643

Table 3.13: T-Test Output for Hypothesis 8

3.4.3.2.9 Hypothesis 9

The null and alternate statements for hypothesis 9 are as follows:

H9o: The four independent variables will not significantly explain the variance in overall usability.

H9a: The four independent variables will significantly explain the variance in overall usability.

As can be examined in Table 3.14 within the model summary table, the R Square is 0.753. This is the amount of variance that is explained in our dependent variable by our four independent variables: functionality, efficiency, interface usability and computer attitude. The F statistic produced (F=292.99) is significant at the 0.000 (rounded) level. Concluding 75 percent of the variance within our dependent variable is significantly explained by our four independent variables. Therefore hypothesis 9 can be rejected.

The table of coefficients shows us that functionality is the most important in explaining the variance within the sample (Beta=0.471). The second important variable is interface usability (Beta=0.412). Both are significant. Efficiency and attitude are not significant predictors.

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	Efficiency, Attitude, Functionality, Interface Usability[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: OverallUsability

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,869[a]	,756	,753	,57832

a. Predictors: (Constant), Efficiency, Attitude, Functionality, InterfaceUsability

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	391,956	4	97,989	292,986	,000[a]
	Residual	126,757	379	,334		
	Total	518,713	383			

a. Predictors: (Constant), Efficiency, Attitude, Functionality, InterfaceUsability

b. Dependent Variable: OverallUsability

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	,243	,163		1,493	,136
	Functionality	,511	,056	,471	9,090	,000
	InterfaceUsability	,420	,055	,412	7,693	,000
	Attitude	-,031	,031	-,032	-,997	,319
	Efficiency	,040	,041	,044	,961	,337

a. Dependent Variable: OverallUsability

Table 3.14: Regression Output for Hypothesis 9

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	Interface Usability, Functionality[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: OverallUsability

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,869[a]	,755	,753	,57798

a. Predictors: (Constant), InterfaceUsability, Functionality

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	391,435	2	195,718	585,873	,000[a]
	Residual	127,278	381	,334		
	Total	518,713	383			

a. Predictors: (Constant), InterfaceUsability, Functionality

b. Dependent Variable: OverallUsability

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	,195	,152		1,281	,201
	Functionality	,516	,053	,476	9,780	,000
	InterfaceUsability	,434	,050	,426	8,763	,000

a. Dependent Variable: OverallUsability

Table 3.15: Regression Output for Functionality and Interface Usability

Further analysis of multiple regression amongst overall usability, functionality and interface usability clearly shows that these two variables are important predictors of perceived usability (Table 3.15). The amount of variance explained is 0.753. The removal of the not significant variables of computer attitude and efficiency has not decreased the amount of variance explained as expected.

3.4.4 Evaluation of Research Results

The conducted field study provided in-depth information in relation to our research objective. Before evaluating the findings, one should be cautious to the context dependent nature of the “usability” concept. Thus these results are only valid under the specific research context. Any comparison of these results to other very different contexts is not possible, at all. However, further studies based on psychometric evaluation of the developed questionnaire within similar contexts should be encouraged.

It has been proved that “computer attitude” is an important variable that impacts “perceived usability”. Also it was demonstrated that this variable is decisive in one’s evaluation of “perceived functionality”, “interface usability” and “time-efficiency” concepts. In e-commerce scenarios, “computer attitude” variable has been found to be an important contributor regarding to potential user’s evaluation of the specific site. However, computer attitude has not been found to be a valid predictor of perceived usability thus it is not possible to predict one’s perception of a specific e-commerce site to a degree, just by examining his / her computer attitude. The high correlations amongst study independent and dependent variables requires further investigation and optimization of ISO 9126-1 (2000) usability model for internet environment. Potential examination points are; definitions of functionality and interface usability, development and optimization of specific questionnaire items for after-scenario testing purposes which measure the aspects that the new and optimized definitions require and examination of promising new variables.

The multiple regression model has been successful in explaining 0.76 percent of the variance within perceived usability. Functionality and interface usability has been found to be significant predictors whereas computer attitude and time-efficiency has been found to be not significant.

One interesting finding has been the impact of “gender” upon computer attitude. Contrary to expectations, there have been no significant differences found regarding “computer attitude” amongst males and females within university students in Istanbul. This is in contrast with the findings of previous studies conducted in United States and United Kingdom but, in harmony with similar studies based in ex-communist countries. However, we must note that the population definitions for these countries show differences. It would be interesting to compare this finding to a similar population definition in other countries. Actually speaking this finding is normal as university students are currently using computers in almost every part of their lives. It is possible that the findings will not be the same if one is to examine a different population within Turkey such as “internet users within Turkey”.

As expected there have been significant differences between students who are studying towards their degrees in technical departments and non-technical departments in their computer attitude. Generally speaking students who are studying in technical departments tend to have more positive computer attitude than students who are studying in non-technical departments.

3.4.5 Implications for Further Research

The results showed high correlations (0.7-0.8) amongst study variables. This has been the main limitation of the study. Careful examination of the high-correlations suggested that the items of CSUQ are not suitable for the research scenario context. For example, users can highly correlate an item which is measuring functionality with overall perceived usability. This has been a frequent feedback that has been returned by interviewers. Therefore, focusing into the definitions of these study variables for an e-commerce context and redeveloping the questionnaire with new items is a further research objective. Further analysis regarding the model’s ability in predicting perceived usability should be done by utilizing new questionnaire designs.

There is still much to do if one is to develop a model which is optimized to a specific context upon ISO 9126-1 (2000) model. The developed model is an ideal start point for the specific context of this scenario, however much research and careful interpretation regarding to the model variables optimization has to be done. The measurement tools which will be developed for the research context should also reflect this view. As a result for almost any different context, there is much work to be done if one is to optimize a model upon ISO 9126-1 (2000).

Cross comparisons of the findings amongst different populations should also be applied for future research designs. This is especially important for evaluation of high correlations amongst study variables found in this research's population. If one is to find similar conditions in larger populations which deviate more in study variables then the high correlations found within this study is an exception for this population. Otherwise, the redefinition of the variables for this context should be revised with further research as indicated in the above paragraph.

Comparative analysis of findings amongst different societies, countries and cultures should be also explored. This will supply in-depth information regarding to the model's ability and optimization for similar research scenarios under different cultural and socio-economical contexts.

Conclusion

Optimization of the ISO 9126-1 (2000) model into a specific context driven scenario and testing the developed questionnaire upon this model in a field study was conducted. The population was the university students in Istanbul. The major objective of the research was to examine the impact of computer attitude upon perceived usability.

It has been proved that computer attitude impacts overall usability perception in the scenario of the research. Therefore attitude can be considered to be an important variable when one's objective is to examine a set of users' perceived usability in a similar context. If the current level of attitude is not favorable / negative towards a particular system, in a similar context, then it is highly likely that the overall usability perceived will also be lower. The practical implication and important finding of the research is that professionals have to be aware of this fact when they are to evaluate a similar context. In the existence of such users with negative attitudes towards computers, the results found can be subjective rather then being objective and thus misrepresenting the current objective usability levels of the system in evaluation. In such a case it is suggested that such users are isolated from the original study for preventing their subjective evaluations which can result in misleading conclusions. Users' attitude towards computers also impact other research variables studied; functionality, interface usability and efficiency to a high degree.

Of the 9 hypothesis tested, 7 were substantiated. Of these 7, four that were substantiated was regarding computer attitude and its impact upon other research variables. The fifth hypothesis was about the model's capability in satisfactorily explaining the variance within perceived usability. The developed model was successful in explaining 75 percent of the variance within the dependent variable, perceived usability. Contrary to expectations, males did not have significant differences in attitudes towards computers then females. Computer attitude of students did differ by technical versus non-technical departments. From the results of the multiple regression analysis functionality and interface usability were found to be important variables affecting perceived usability. Out of the four study variables only these two were found to be significant within the multiple regression analysis. The remaining variables; attitude and efficiency were not found to be significant.

The means of research variables were all high for the selected e-commerce site within the scenario representing high usability levels perceived. Especially speaking, computer attitude amongst university students in Istanbul was high with a low standard deviation. Thus one can assume that the university students' attitudes towards computers are highly favorable.

Regarding the developed questionnaire and its potential use for an evaluation tool for e-commerce systems, there is still much to do. The high correlations within the study variables require the careful examination of the variables for this context for ensuring and improving the validity. Focusing onto functionality and interface usability are ideal start points. These two variables were successful in satisfactorily explaining 75 percent of the variance within perceived usability. However, it is advised that further research takes into account different populations. Conducting a research within a population who is very high on a study variable / variables makes it difficult and sometimes impossible to collect / analyze a distributed data pattern which can supply more information. Still, the questionnaire's current state is satisfactory in being able to predict perceived usability levels in similar contexts.

References

Alreck, P.L., Settle, R.B. (1994).The Survey Research Handbook. Irwin Professional Publishing, Chicago, IL

Balka, E., Smith, R. (Eds.), (2000), Women Work and Computerization, Boston: Kluwer

Bandura, A. (1997), The Exercise of Control. New York: W. H. Freeman

Bandura, A. (1982), Self-efficacy Mechanism in Human Agency. American Psychologist, 37(2) 122-147

Bear, G. G., Richards, H. C., Lancester, P. (1987), Attitudes Toward Computers: Validation of a Computer Scale. Journal of Educational Computing Research, 3, 207–218

Beckers, J., Schmidt, H. (2001), The Structure of Computer Anxiety: A Six Factor Model. Computers in Human Behavior, 17(1), 35–49

Bessie're Katie, 1, Newhagen John E.,*, Robinson John P., Schneiderman Ben (2006), A Model for Computer Frustration: The Role of Instrumental and Dispositional Factors on Incident, Session, and Post-session Frustration and Mood, Computers in Human Behavior 22 (2006) 941–961

Bevan, N. (1995a), Measuring Usability as Quality of Use, Journal of Software Quality (4), p. 115-140

Bevan, N. (1995b), Usability is Quality of Use, In Anzai, Ogawa, Proceedings of the 6th International Conference on Human Computer Interaction, July 1995.

Bevan, N. (2001), International Standards for HCI and Usability, International Journal of Human-Computer Studies 55 (4), p. 533-552

Brosnan, M. J. (1998), The Impact of Computer Anxiety and Self-efficacy upon Performance, Journal of Computer Assisted Learning, 14, 223–234

Bozionelos, N. (2001), Computer Anxiety: Relationship with Computer Experience and Prevalence. Computers in Human Behavior, 17, 213–224

Brosnan, M., Lee, W. (1998), A Cross-cultural Comparison of Gender Differences in Computer Attitudes and Anxieties: The United Kingdom and Hong Kong. Computers in Human Behavior, 14, 559–577

Cambre, M. A., Cook, D. L. (1985), Computer anxiety: Definition, Measurement, and Correlates. Journal of Educational Computing Research, 1(1), 37–54

Chin, J.P., Diehl, V.A., Norman, K.L. (1988), Development of an Instrument Measuring User Satisfaction of the Human–computer Interface, In: Proceedings of the CHI Conference on Human Factors in Computing Systems. ACM Press, New York, pp.213 –218

Cohen, B. A., Waugh, G. W. (1989), Assessing Computer Anxiety. Psychological Reports, 65(1), 735–738

Colley, A., Gale, M., Harris, T. (1994), Effects of Gender Role Identity and Experience on Computer Attitude Components, Journal of Educational Computing Research, 10(2), 129–137

Constantine, L.L., Lockwood, L.A.D. (1999), Software for Use: A Practical Guide to the Models and Methods of Usage-Centered Design. Addison-Wesley, New York, NY

Dambrot, F. H., Watkins-Malek, M. A., Silling, S. M., Marshall, R. S., Garver, J. A. (1985), Correlates of Sex Differences in Attitudes Toward and Involvement with Computers. Journal of Vocational Behavior, 27, 71–86

Davis, F. D. (1986). A Technology Acceptance model for Empirically Testing New End-user Information Systems: Theory and Results, Unpublished Doctoral Dissertation, Sloan School of Management, Massachusetts Institute Technology, Cambridge, MA

Davis, F.D. (1989), Perceived Usefulness, Perceived Ease of Use and User Acceptance of Information Technology, MIS Quarterly 13 (3), 319 –340

Doll, W.J., Xia, W., Torkzadeh, G.A (1994), Confirmatory Factor Analysis of the End-user Computing Satisfaction Instrument, MIS Quarterly 18 (4), 453 –461

Durndell A, Haag Z. (2002), Computer Self Efficacy, Computer Anxiety, Attitudes towards the Internet and Reported Experience with the Internet, By Gender, In an East European Sample, Computers in Human Behavior 18 (2002) 521–535

Durndell, A., Thomson, K. (1997), Gender and Computing: A Decade of Change? Computers and Education, 28(1), 1–9

Fishbein, M., Ajzen, I. (1975), Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research. Boston, MA: Addison-Wesley

Folmer, E., Bosch, J. (2004), Architecting for Usability, The Journal of Systems and Software, 2004, 70, p61-78

Gardner, D. G., Dukes, R. L., Discenza, R. (1993). Computer Use, Self-Confidence, and Attitudes: A causal Analysis, Computers in Human Behavior, 9, 427–440

Garland, Kate J., Noyes, Jan M. (2004) Computer Experience: A Poor Predictor of Computer Attitudes. Computers in Human Behavior 20, 823–840

Gediga Günther, Hamborg Kai-Christoph, Düntsch, Ivo (?), Evaluation of Software Systems

Geissler, J. E., Horridge, P. (1993), University Students Computer Knowledge and Commitment to Learning, Journal of Research on Computing in Education, 25, 347–365

Hackett, G., Betz, N. E. (1989), An Exploration of the Mathematics Self-efficacy / Mathematics Performance Correspondence, Journal for Research in Mathematics Education, 20, (?) 261-273

Hartson, H.R., Castillo, J.C. Kelso, J. Kamler, J., Neale, W.C. (1996). Remote Evaluation: The Network as an Extension of the Usability Laboratory. In: Proceedings of CHI'96 Human Factors in Computing Systems. Pp. 228-235

Hix, D., Hartson, H.R. (1993), Developing User Interfaces: Ensuring Usability through Product and Process. Wiley and Sons, New York, NY (chapter 2)

Igbaria, M., Chakrabarti, A. (1990), Computer Anxiety and Attitudes towards Microcomputer Use, Behavior and Information Technology, 9(3), 229–241

ISO 9126 (1991), Software Production Evaluations-Quality Characteristics and Guidelines for their Use, ISO DIS 9126

ISO 9126-1 (2000), Software Engineering – Product Quality – Part -1: Quality Model

ISO 9241-11 (1994), Ergonomic Requirements for Office Work with Visual display Terminals (VDTs), Part-11 Guidance on Usability

Jones, T., Clark, V. A. (1995), Diversity as a Determinant of Attitudes: A Possible Explanation of the Apparent Advantage of Single-sex Settings. Journal of Educational Computing Research, 12, 51–64

Kay, R. H. (1989), A Practical and Theoretical Approach to Assessing Computer Attitudes The Computer Attitude Measure (CAM), Journal of Research on Computing in Education, 21, 456–463

Kay, R. H. (1993). An exploration of Theoretical and Practical Foundations for Assessing Attitudes towards Computers: The Computer Attitude Measure (CAM). Computers in Human Behavior, 9, 371–386

Kurosu, M., Kashimura, K. (1995), Determinants of the Apparent Usability, In: Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (Vol. 2, pp. 1509–1514). New York: IEEE

Levine, T., Donitsa-Schmidt, S. (1998), Computer Use, Confidence, Attitudes and Knowledge: A Causal Analysis. Computers in Human Behavior, 14, 125–146

Lewis, J.R. (1992a), Psychometric Evaluation of the Post-study System Usability Questionnaire: the PSSUQ. In Proceedings of the Human Factors Society 36th Annual Meeting, HFES, Santa Monica, CA, pp.1259 –1263

Lewis, J. R. (1992b), Psychometric Evaluation of the Computer System Usability Questionnaire: The CSUQ (Tech. Report 54.723), Boca Raton, FL: International Business Machines Corporation

Lewis, J.R. (1995), IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use/James R. Lewis. International Journal of Human –Computer Interaction 7 (1), 57 –78, pp.375 –380

Liaw, S. (2002), An Internet Survey for Perceptions of Computers and the World Wide Web: Relationship Prediction and Difference. Computers in Human Behavior, 18, 17–35

Loyd, B. H., Gressard, C. (1984), Reliability and Factorial Validity of Computer Attitude Scales, Educational and Psychological Measurement, 44(2), 501–505

Macleod, M. (1994), Usability: Practical Methods for Testing and Improvement. In: Proceedings of the Norwegian Computer Society Software '94 Conference, Sandvika, Norway

Maurer, M. (1994), Computer Anxiety Correlates and What They Tell Us: A Literature Review. Computers in Human Behavior, 10(3), 369–376

Molich, R., Nielsen, J. (1990). Improving a Human-computer Dialogue, Communications of the ACM 33, 3 (March), 338-348

Nielsen, J. (1993a). Usability Engineering, Academic Press, San Diego, CA

Nielsen, J. (1993b), IEEE Computer Vol. 26, No. 11 (November 1993), pp. 32-41

Nielsen, J. (1994a), Enhancing the Explanatory Power of Usability Heuristics, Proc. ACM CHI'94 Conf. (Boston, MA, April 24-28), 152-158

Nielsen, J. (1994b), Heuristic Evaluation, In Nielsen, J., MACK, R.L., Usability Inspection Methods, John Wiley and Sons, New York, NY

Potosky, D., Bobko, P. (1998). The Computer Understanding and Experience Scale: A Self-Report Measure of Computer Experience. Computers in Human Behavior, 14, 337–348

Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., Carey, T. (1994), Human-Computer Interaction, Addison-Wesley, Reading, MA

Reinen, I., Plomp, T. (1997), Information Technology and Gender Equality: A Contradiction in Terminis? Computers and Education, 28(2), 65–78

Rosen, L. D., Sears, D. C., Weil, M. M. (1987), Computerphobia, Behavior Research Methods, Instruments, and Computers, 19, 167–179

Schakel, B. (1991), Designing the User Interface: Strategies for Effective Human-Computer Interaction, Addison-Wesley, MA

Sekaran, U. (2000), Research Methods for Business Third Edition, John Wiley & Sons, 295.

Shashaani, L. (1994), Gender-Differences in Computer Experience and Its Influence on Computer Attitudes, Journal of Educational Computing Research, 11(4), 347–367

Shen Demei, Laffey James, Lin Yimei, Huang Xinxin (2006), Social Influence for Perceived Usefulness and Ease-of-Use of Course Delivery Systems, Journal of Interactive Online Learning, Colume 5, Number 3, winter 2006, ISSN: 1541-4914

Shneiderman, B. (1986), Designing the User Interface: Strategies for Effective Human-Computer Interaction. Addison-Wesley, Reading, MA

Smith, B., Caputi, P., Crittenden, N., Jayasuriya, R., & Rawstorne, P. (1999), A Review of the Construct of Computer Experience. Computers in Human Behavior, 15, 227–242

Smith, B., Caputi, P., Rawstorne, P. (2000), Differentiating Computer Experience and Attitudes toward Computers: An Empirical Investigation. Computers in Human Behavior, 16, 59–81

Smith, M. Sheila (Ball State University) (?), An Examination of the Computer Self-Efficacy and Computer-Related Task Performance Relationship, sweber2@bsu.edu

Soken, N. Reinhart, B., Vora, P., Metz, S. (1993), Methods for Evaluating Usability (Section5B), Honeywell, Dec. 1993

Torkzadeh, G., Angulo, I. E. (1992), The Concepts and Correlates of Computer Anxiety. Behavior and Information Technology, 11(1), 99–108

Torkzadeh, G., Koufteros, X. (1994), Factoral Validity of a Computer Self-Efficacy Scale and the Impact of Computer Training, Educational and Psychological Measurement, 54(3), 813–921

Venkatesh, V., Davis, F. D. (2000), A Theoretical Extension of the Technology Acceptance Model: Four longitudinal field studies. Management Science, 46, 186-204

Vora, P., Helander, M. (1995).A Teaching Method as an Alternative to the Concurrent Think-Aloud Method for Usability Testing. In Anzai, Y., Ogawa, K., Mori, H.(Eds.), Symbiosis of Human and Artifact, pp.375 –380

Ward, R.D., Marsden (2003), P.H. Physiological Responses to Different Webpage Designs, Int. Journal of Human-Computer Studies 59 (2003)199 –212

Wharton, C., Rieman, J., Lewis, C., Polson, P. (1994). The Cognitive Walkthrough: a Practitioner’s Guide. In: Nielsen, J., Mack, R.L. (Eds.), Usability Inspection Methods. John Wiley and Sons, New York, NY

William J. Doll, Anthony Hendrickson, Xiaodong Deng (1998) Using Davis's Perceived Usefulness and Ease-of-use Instruments for Decision Making: A Confirmatory and Multigroup Invariance Analysis
Decision Sciences 29 (4), 839–869 doi:10.1111/j.1540-5915.1998.tb00879.x

Whitely, B. (1997), Gender Differences in Computer Related Attitudes and Behavior: a Meta analysis. Computers in Human Behavior, 13(1), 1–22

Wixon, D., Jones, S., Tse, L., Casaday, G. (1994). Inspections and Design Reviews: Framework, History and Reflection, In: Nielsen, J., Mack, R.L. (Eds.), Usability Inspection Methods. John Wiley and Sons, New York, NY, pp.77 –103

Wixon, D., Wilson, C. (1997), The Usability Engineering Framework for Product Design and Evaluation, In Helander, M.G. et al. (Eds), Handbook of Human-Computer Interaction, Elsevier, North-Holland, pp. 653-688

Wright, R. (1997), Women in Computing: A Cross National Analysis. In R. Lander, & A. Adam (Eds.), Women in Computing (pp. 72–83), Exeter, England: Intellect Books

Zhang, Z., Basili, V., Shneiderman, B. (1998), An Empirical Study of Perspective-Based Usability Inspection. In: Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting, Chicago.pp.1346 –1350

Zhang, Z. (2001), Overview of Usability Evaluation Methods, http://www.cs.umd.edu/~zzj/UsabilityHome.html , 01-01-2008.

Appendix A Developed Questionnaire

PaketTicaret.com Elektronik Ticaret Portalı Değerlendirme Anketi

Lütfen tüm soruları cevaplayınız.

1. Bilgisayarlar ile çalışmayı seviyorum.

Kesinlikle Katılmıyorum (1)	1	2	3	4	5	6	7	(7) Kesinlikle Katılıyorum
	○	○	○	○	○	○	○	

2. Genel olarak, bu e-ticaret sitesini kullanmanın basitliği beni tatmin etti.

Kesinlikle Katılmıyorum (1)	1	2	3	4	5	6	7	(7) Kesinlikle Katılıyorum
	○	○	○	○	○	○	○	

3. Bana verilen tüm görevleri, bu e-ticaret sitesinde, etkin bir şekilde tamamlayabildim.

Kesinlikle Katılmıyorum (1)	1	2	3	4	5	6	7	(7) Kesinlikle Katılıyorum
	○	○	○	○	○	○	○	

4. Bu e-ticaret sitesini kullanmak basit.

Kesinlikle Katılmıyorum (1)	1	2	3	4	5	6	7	(7) Kesinlikle Katılıyorum
	○	○	○	○	○	○	○	

5. İşimin bilgisayar kullanımını gerektiren konularına hevesle bekliyorum.

Kesinlikle Katılmıyorum (1)	1	2	3	4	5	6	7	(7) Kesinlikle Katılıyorum
	○	○	○	○	○	○	○	

6. Bu e-ticaret sitesinde bana verilen görevleri hızlı bir şekilde tamamladım.

Kesinlikle Katılmıyorum (1) 1 2 3 4 5 6 7 (7) Kesinlikle Katılıyorum

7. Bu e-ticaret sitesi karşılaştığım sorunları çözebilmemi sağlayan açık ve net mesajlar / yönlendirmeler verdi.

Kesinlikle Katılmıyorum (1) 1 2 3 4 5 6 7 (7) Kesinlikle Katılıyorum

8. Bilgisayar kullanırken çabuk sıkılırım.

Kesinlikle Katılmıyorum (1) 1 2 3 4 5 6 7 (7) Kesinlikle Katılıyorum

9. Aradığım bilgiyi bulmak basitti.

Kesinlikle Katılmıyorum (1) 1 2 3 4 5 6 7 (7) Kesinlikle Katılıyorum

10. Bana verilen görevlerimi bu e-ticaret sitesinde verimli bir şekilde tamamlayabildim.

Kesinlikle Katılmıyorum (1) 1 2 3 4 5 6 7 (7) Kesinlikle Katılıyorum

11. E-ticaret sitesindeki bilgiler / içerik verilen görevlerimi tamamlarken bana etkin bir şekilde yardımcı oldu.

Kesinlikle Katılmıyorum (1) 1 2 3 4 5 6 7 (7) Kesinlikle Katılıyorum

12. Bilgilerin ekrandaki düzeni / sunumu açık ve netti.

Kesinlikle Katılmıyorum (1) 1 2 3 4 5 6 7 (7) Kesinlikle Katılıyorum

13. E-ticaret sitesinin genel tasarımını / arayüzünü beğendim.

Kesinlikle Katılmıyorum (1) 1 2 3 4 5 6 7 (7) Kesinlikle Katılıyorum

14. Bir kez bilgisayar ile çalışmaya başlayınca, kendimi bilgisayardan alıkoymaktan zorlanıyorum.

Kesinlikle Katılmıyorum (1) 1 2 3 4 5 6 7 (7) Kesinlikle Katılıyorum

15. Bu e-ticaret sitesi beklediğim tüm fonksiyon ve özelliklere sahip.

Kesinlikle Katılmıyorum (1) 1 2 3 4 5 6 7 (7) Kesinlikle Katılıyorum

16. Bu e-ticaret sitesinin arayüzünü / kullanıcı ekranlarını kullanmayı sevdim.

Kesinlikle Katılmıyorum (1) 1 2 3 4 5 6 7 (7) Kesinlikle Katılıyorum

17. Bilgisayar kullanmak asabımı bozuyor.

Kesinlikle Katılmıyorum (1) 1 2 3 4 5 6 7 (7) Kesinlikle Katılıyorum

18. Genel olarak, bu e-ticaret sitesi beni tatmin etti.

Kesinlikle Katılmıyorum (1) 1 2 3 4 5 6 7 (7) Kesinlikle Katılıyorum

Anket Cevaplarımı İlet >>>