T.C.
İSTANBUL UNIVERSITY
INSTITUTE OF GRADUATE STUDIES IN
SCIENCE AND ENGINEERING

Ph.D. THESIS

A USABILITY EVALUATION METHODOLOGY WITH A QUESTIONNAIRE SUPPORTED WITH QUALITATIVE DATA FOR WEB SITE USER INTERFACE

Mehmet Salih GÖÇERİ

Department of Industrial Engineering

Ph. D. transferred from Fatih University which has been closed

SUPERVISOR
Asst. Prof. Dr. Özgür UYSAL

December, 2016

İSTANBUL

APPROVAL PAGE

This is to certify that I have read this thesis written by Mehmet Salih GÖÇERİ and that in my opinion it is fully adequate, in scope and quality, as a thesis for the degree of Doctor of Philosophy in Industrial Engineering.

Assist. Prof. Özgür UYSAL
Thesis Supervisor

I certify that this thesis satisfies all the requirements as a thesis for the degree of Doctor of Philosophy in Industrial Engineering.

Prof. Mehmet ŞEVKLİ
Head of Department

Examining Committee Members

Prof. Selim ZAİM

Prof. Nizamettin BAYYURT

Assoc. Prof. Mustafa Şeref AKIN

Assoc. Prof. Ali TÜRKYILMAZ

It is approved that this thesis has been written in compliance with the formatting rules laid down by the Graduate School of Sciences and Engineering.

Prof. Nurullah ARSLAN
Director

June 2016

A USABILITY EVALUATION METHODOLOGY WITH A QUESTIONNAIRE SUPPORTED WITH QUALITATIVE DATA FOR WEB SITE USER INTERFACE

Mehmet Salih GÖÇERİ

Ph.D. Thesis Industrial Engineering
June, 2016

Thesis Supervisor: Asist. Prof. Özgür UYSAL

ABSTRACT

In this paper, a unique methodology for usability evaluation is proposed. The underlying factors for usability have been determined with exploratory and confirmatory factor analysis. The results of the analysis produced a valid and reliable questionnaire for user interface (QUIN) with eight different dimensions. Additionally, qualitative data from participants have been collected at the end of each survey. Afterwards, the result of quantitative data from questionnaire responses and qualitative findings from user feedback are compared. The methodology is designed to be used for any web site, especially for informational web sites. A case study is performed for the student portal of a Turkish university in order to validate the methodology. QUIN has identified the most critical usability problems supported with user feedback.

Keywords: Usability Evaluation, Human Computer Interaction, User Interface, Web Site Quality, Questionnaire, Factor Analysis.

WEB SİTESİ KULLANICI ARAYÜZÜNÜN NİTEL VERİ İLE DESTEKLİ ANKET YÖNTEMİYLE KULLANILABİLİRLİĞİNİ DEĞERLENDİRME METODOLOJİSİ

Mehmet Salih GÖÇERİ

Doktora Tezi – Endüstri Mühendisliği
Haziran, 2016

Danışman: Yrd. Doç. Dr. Özgür UYSAL

ÖZ

Bu çalışmada kullanılabilirlik değerlendirmesi için yeni bir yöntem sunulmuştur. Kullanılabilirliğin asıl faktörlerini belirlemek için keşfedici ve doğrulayıcı faktör analizi uygulanmıştır. Analiz sonuçları kullanıcı arayüzü için 8 faktörden oluşan geçerli ve tutarlı bir anket ortaya koymuştur. Ayrıca katılımcılardan her anketin sonunda nitel veri yani kullanıcı görüşleri toplanmış olup, bu veriler ile anket cevaplarından gelen nicel veri ile karşılaştırılmıştır. Yöntem herhangi bir sayfa için uygulanabilir olup özellikle bilgi amaçlı web siteleri için tasarlanmıştır. Vaka analizi Türkiye'de özel bir üniversitenin arayüzüne uygulanmış, elden edilen veriler ile yöntemin geçerliliği test edilmiştir. Sunulan metodoloji en kritik kullanılabilirlik problemlerini kullanıcı görüşleri ile birlikte belirlemiştir.

Anahtar Kelimeler: Kullanılabilirlik Değerlendirmesi, Bilgisayar İnsan Etkileşimi, Kullanıcı Arayüzü, Web Site Kalitesi, Anket, Faktör Analizi.

To my family,

ACKNOWLEDGEMENT

I express sincere appreciation to Assist. Prof. Özgür UYSAL and Prof. Selim Zaim for their guidance, insight and motivation throughout the research.

Thanks go to the other faculty members, Assoc. Prof. Dr. Mustafa Şeref Akın and Assoc. Prof. Ali Türkyılmaz, for their valuable suggestions and comments. Additionally, the technical assistance of Prof. Nizamettin Bayyurt is gratefully acknowledged. Moreover, motivation and valuable suggestion of Prof. Mehmet Şevkli will be always remembered.

Finally, I express my thanks and appreciations to my family for their understanding, motivation, and patience during the thesis process.

TABLE OF CONTENTS

LIST OF TABLES

TABLE

LIST OF FIGURES

FIGURE

LIST OF SYMBOLS AND ABBREVIATIONS

SYMBOL/ABBREVIATION

SPSS	Statistical Package for the Social Science
AMOS	Analysis of Moment Structures
QUIN	Questionnaire for User Interface
HE	Heuristic Evaluation
EFA	Exploratory Factor Analysis
CFA	Confirmatory Factor Analysis

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

Internet usage for information sharing, data transferring, personal servicing, and entertaining has been incredibly increased over the last decades (Navimipour, 2015). These services are provided to the users through the web sites. Nowadays, the web sites are more than a simple and single entity; namely, they are not only used for information search and entertainment (Navimipour and Zareie, 2015). Web sites can be categorized into four groups; entertainment, information, communication, and commerce (Lee and Koubek, 2010).

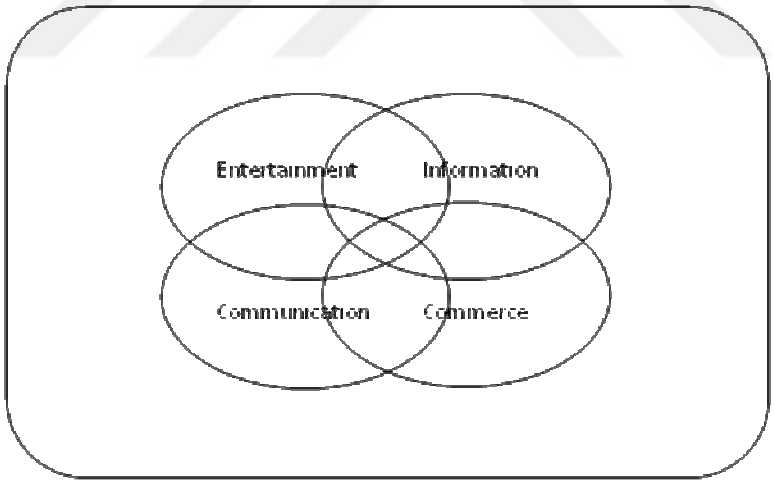


Figure 1.1 Categories of the web sites (Lee and Koubek, 2010).

The history of web site has started over couple of decades. First web sites were built by large companies, which have some standards such as putting their logos in each page. These web sites were expensive to operate and using a large database of companies that were causing slowness in their speed. Later, some web sites appeared in the market, which navigates fast but without standards or logos. Due to lack of some standards, these web pages were usually unattractive to get people attention and they could not exist longer in online world. Therefore, companies spent an incredible effort

in order to expand their businesses using internet in their daily activities over last decades. Designing the good web sites reflect a strong relationship with customers as well as a reputation for the companies (Chiagouris and Wansley, 2000). Additionally, getting the attention of people has become a crucial goal for those companies to improve the standard of web sites so that they can fulfill customers' expectations and increase their revenues in online world (Iwaardena et al, 2003).

However, most of the web sites still contain many problems (Becker and Mottay, 2001, Chau and Wong, 2010, Treiblmaier and Pinterits, 2010, Tung et al., 2009). The common problems can be defined as difficulties of understanding the content, difficulties in navigation, disorientation, lack of customization, reliability, consistency on format, efficiency on search capability, flexibility of system help function, content update, security, speed, responsiveness and so on (Downing and Liu, 2011; Fogli and Guida, 2014). These problems are regarding the quality and usability of web sites.

1.2 USER INTERFACE

Users are the most important target for a web site and their quality perceived by the users is one of the most important factors for their success. In online world, an effective web site design is an essential for an organization to be accepted successful (Muyllea et al., 2003). According to National Science Foundation Graphical (NSFG, 1999) user interface has been recognized as one of the deciding factor that makes electronic commerce possible.

Since there are millions web sites in online world, the way how they operate when providing services to their users are the same. In other words, all companies or organization even individual service providers get in touch with their customers/user over an interface of web site. An online organization interacts with current or potential customers through the interface of the web sites. Therefore, the most powerful aspect of organization for delivering services to customers has become the interface of web sites.

The interface plays as web sites or web portals and users and service providers interacts with each other through the interface of computers (Muyllea et al., 2003) tablets, smart phones or some devices. Users will connect to the companies through the interface of web sites anywhere around the world. User interface of web sites are the most important way to keep customers connected to companies. Similarly, if the service

provider as university connects to students through the interface of university web site where registration information, recent news, grades, and some other information is provided to students.

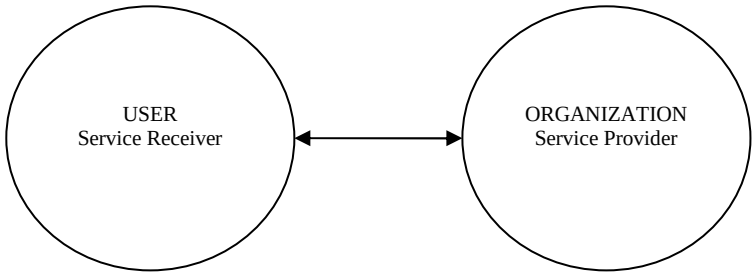


Figure 1.2 Interactions between user and service provider.

Despite the fact that the purpose of web sites is different, the way how services are delivered to customer through web sites and the way how users connected to service providers are the same. In other words, either each web sites users or providers will interact to other side with an interface.

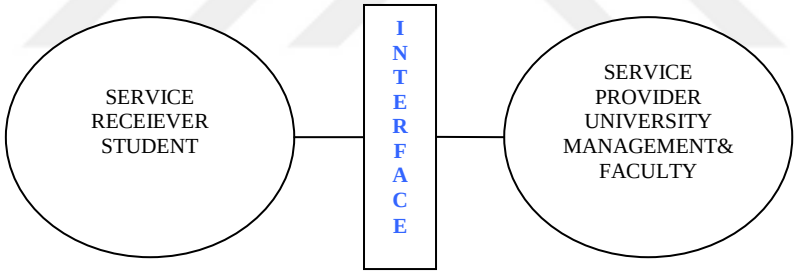


Figure 1.3 Student gets service from both university management and faculty.

As seen above, the interaction between university and faculty with student occurs as shown in Figure 1.3. Students get some service from university and their professors through university web portal.

While either users or service providers use the interface, they also experience a quality of the service received. Both users and service providers define the meaning of the quality in accordance with their perception. How quality can be defined for all? Since students get a service from the interface, the quality of the service needs to be considered.

1.3 IMPORTANCE OF THE USABILITY RESEARCH

Developing a usable, highly qualified, well-designed web site is one of the most complicated issues in online world. The user interfaces of the most web sites contain many problems in terms of usability (Becker and Mottay, 2001, Chau and Wong, 2010, Treiblmaier and Pinterits, 2010, Tung et al., 2009). If customers are satisfied with service experienced, they visit the web site again. Dissatisfaction of the web site, due to poor interface, could result an unacceptable learning time of the content, too many clicks to navigate, and longer time of the web pages loading. If users have been offered more usable web sites by competitors, they might visit those web sites (Ratner, 2003). According to Nielsen, if customers cannot find what they look for, then they will not buy it (Nielsen, 2000, “usability rules the web”)

A research has found that if customers cannot find what they are looking for, potential sales will approximately drop by %50 or more. Additionally, if customers had a negative experience when they visit web site at first time, %42 of them do not want to return the web site again (Manning, 1998). In addition, Becker and Mottay expresses that users are running away from unusable web sites, which causes online business failures.

In Another study Jared Spool found that "What percentage of the time are visitors successful at achieving their goals on the best-designed web sites?", and the answer was only 42%. This result could take away dissatisfied customer from even best designed web sites (Spool, 1999). If these websites are Apple, Amazon, and EBay, online market giants, these dissatisfied customer rate could cause billion dollars loss on sales.

Usability of web sites is sometimes about customer perception for the service experienced. Ben Sheiderman found that if web site change designs of interface such a color, position of elements, terminology of button, this could decrease the performance of the web site up to 25%.

Majority of the web site sites violated the simple design principles. There are some evidences regarding how usability affects users’ behaviors as follows:

1. 15 commercial web sites analyzed; even though users the test from correct home page, only 42% of the information was found by users.

2. 62% of web shoppers gave up looking for the item they tried to purchase online.

3. 51% compliance were with simple web usability principles such as "is the site organized by user goals?" and "does a search list retrievals in order of relevance?"

What do they do when they could not find what they are looking for or what if they are not satisfied with the service they took from a web site? They would rethink whether or not it's right place, “Google it”, or visit the competitors web sites.

Before the usability concept, designer developing the web sites without thinking if there could be a problem for users. Spool (2009) claimed that after usability testing, there are many usability problems have been found and fixed which increased satisfaction of users and sales. The issues sometimes were just only removing some disturbing process for registration or changing the whole design of the web site but only adding search button or improving some section of web sites. For example, IBM reports that after redesigning of the web site, usage of search engine in the web site has been decreased %84 by users because the most used feature of web site was search engine. After redesigning the web site, the sales were increased by 400% because users were finding easily what they were looking for without using the search tool as they did before (Tedeschi, 1999).

Users want the web sites to be more usable, easy to use and understand the content quickly. In addition, they are interested some special features such as download speed, trust, responsiveness, and empathy (Downing and Liu, 2011).

As seen below in Figure 1.5, usability studies are becoming more popular compared to other related research fields since 1997. This finding also shows why usability studies are important.

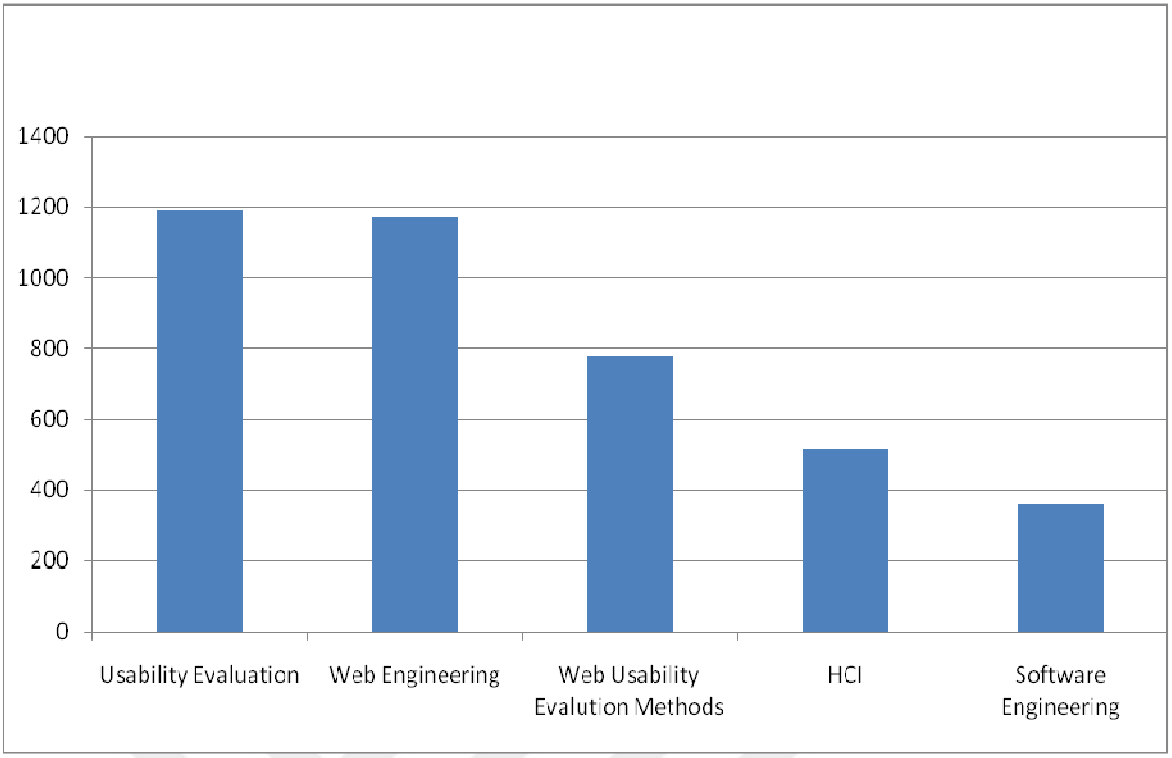


Figure 1.4 Increase percentage of usability studies compared to related research fields (Fernandez et. al., 2011).

1.4 PURPOSE OF THE RESEARCH

There are many approaches developed to assess the quality of web site and their usability. The methods to evaluate the web site usability also have been proposed differently due to lack of consensus on either factors or evaluation methods of usability (Oztekin et al, 2009, Fogli and Guida, 2014). Additionally, usability evaluations methods depend on the purpose of the evaluation, limitation of resources, and preference of researchers (Fernandez et al., 2011). Furthermore, many proposed research methods have either reliability or validity issues (Muyllea et al, 2003, Oztekin et al, 2009, Elling et al., 2012). Moreover, some of usability evaluation methods proposed, especially questionnaires may need to be supported with another evaluation method which detects more usability problems (Hertzum and Jacobsen, 2003, Walji et al, 2014). In this context, the purpose of the research was to construct a unique and comprehensive approach for measuring usability which determines usability factors as well is still needed.

In this study, a methodology for the evaluation of user interface (QUIN) will be designed. Firstly, quantitative data from questionnaire response will be collected. Secondly, a qualitative data from participant at the end of each questionnaire will be collected. Finally, the results of both data collection will be compared to explore the critical usability problems for users because *Hallahan (2001) expressed that "most usability tests rely upon triangulation, which is combining several data gathering techniques—including quantitative and qualitative measures"(p. 226).* The methodology is designed to be used for any web site, especially for informational web site. A case study will be performed for the student portal of a Turkish university in order to validate the methodology. Supporting the survey with user feedbacks could increase the performance of the methods to discover the usability problems. Additionally, comparing the result of survey response with user feedback could provide a superior approach to detect usability problems for informational web site.

The goal of study therefore was to contribute a new approach to literature so that practitioners and scholars can utilize while designing a web site, which put user's perspective into design process with quantitative data. Steps in the research were:

1. Introducing and defining a web site usability evaluation method/checklist;
2. Empirically validating the construct;
3. Identifying the underlying dimensions of the usability checklist and sub factors;

4. Comparing findings with users' feedbacks;

5. Determining challenges of the users while they navigate the web sites with respect to questionnaire and feedbacks;

6. Defining the problems and offering solutions to these problems;

7. Drawing a conclusion and giving recommendations for the future researches.

1.5 RESEARCH QUESTION

In this study, a unique web site measurement checklist will be constructed for web site interface especially for informational web sites such as governmental and

informational web sites. After constructing the checklist, it will be applied to an interface so that validity and reliability of the methodology will be checked. Therefore, research question is “how can I develop a unique web site usability evaluation method so that the most critical factors that affect the users while using the web site interface can be detected”.

After determination of the most critical problems for the web site, the problems will be submitted to the web site designer team and executive of the organization. “If” the development team and executives fix the critical usability problems, web site will be more usable for the users.

1.6 THESIS STRUCTURE

In Chapter 2 is regarding literature review for the terms used in this thesis. The history and definition of each term used in this study were explained.

In Chapter 3 is regarding the construction of instrument. The process of development for the methodology was explained systematically.

In Chapter 4 is regarding data analysis. In addition, reliability and validity of the instrument was checked.

In Chapter 5, the results of qualitative data and quantitative data were compared and the most important usability problems were determined.

In Chapter 6, the research was concluded with findings and future research alternatives were mentioned.

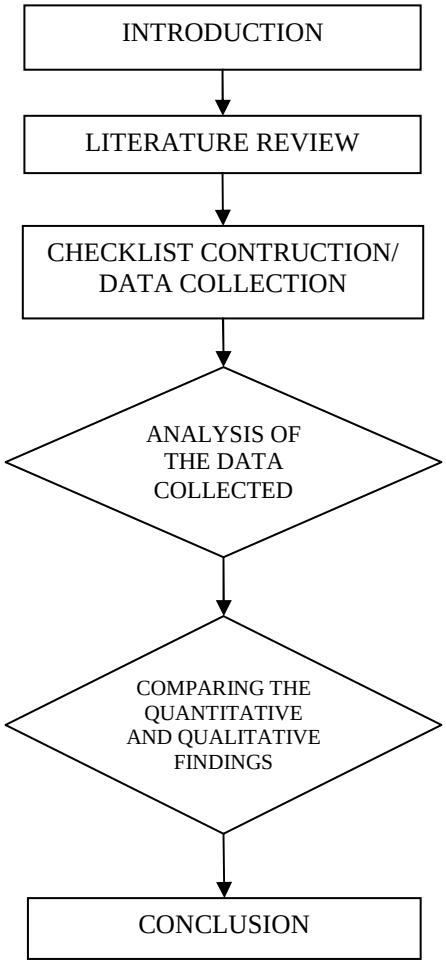


Figure 1.5 Thesis structure

CHAPTER 2

LITERATURE REVIEW

2.1 TODAY IS NOT USABLE FOR TOMORROW

Usability term reflects a process of development, which explained differently over the years. Although, first generation computers, button-based phones and web sites were usable in the past, they are not usable nowadays. Their places have been taken by smart computers and phones, touch-screen interfaces, more effective and efficient devices. However, today's technology, products, interfaces may not be usable for the next generations because today requirements is not enough and usable for tomorrow's expectations.

2.2 LITERATURE REVIEW

In this chapter, the terms related to this study such as history of quality, service industry, Human Computer Interface and usability and web site usability will be reviewed. While focusing on these concepts, the relationship between them will be determined as well. In order to understand the "meaning of usability" better, firstly the history of quality will be reviewed. Afterwards, service quality and Human Computer Interface, usability and web site usability will be explained respectively.

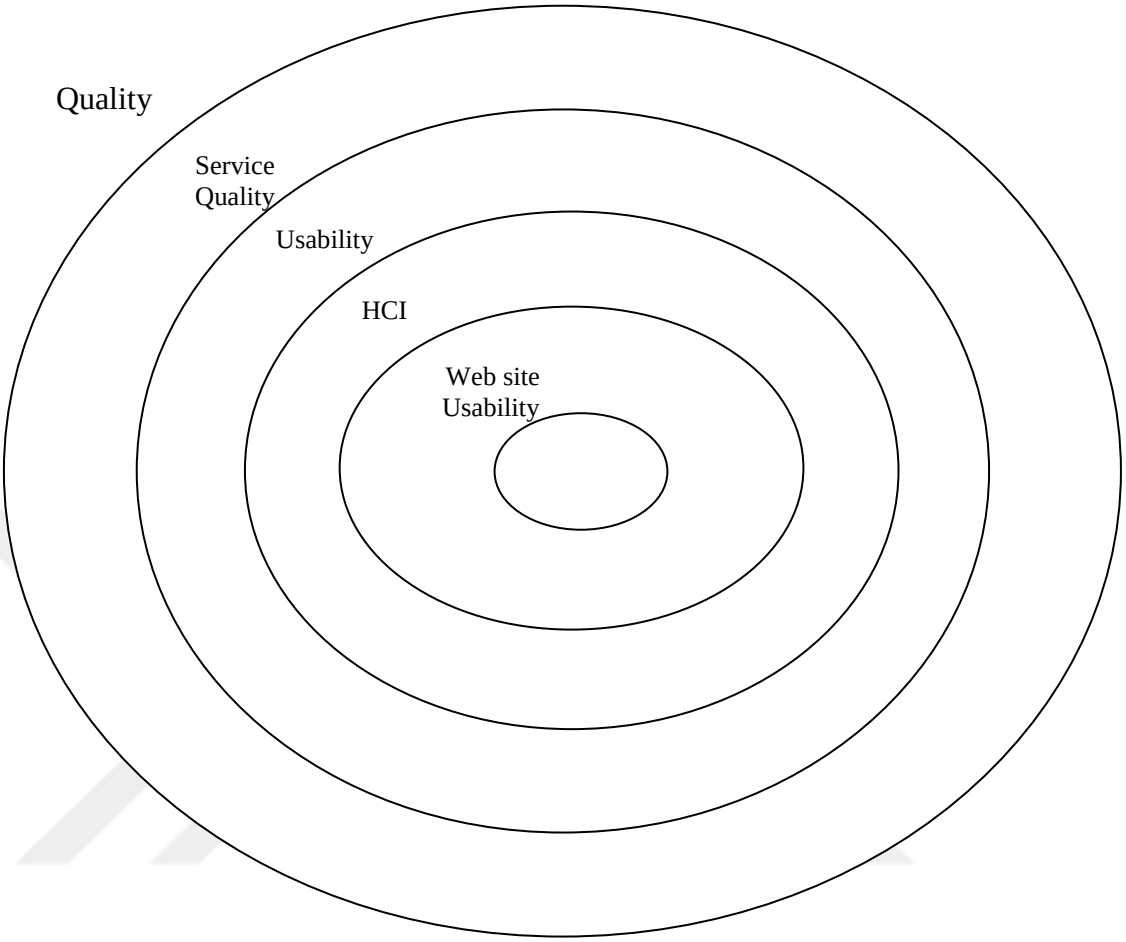


Figure 2.1 Relation between quality, usability, and related areas.

2.2.1 An Overview for Quality Movement

History of quality started as product quality movement later as the evolution from quality control to quality improvement and finally as quality assurance.

Walter Shewhart is considered as ‘grandfather’ of quality control. He designed the first control chart; a statistical processes control and a quality improvement program, and defined his findings as ‘quality assurance’. His idea, reducing variability, was actually quality improvement. His approach to quality control was a process-oriented approach, during the process, using statistics to understand and manage the possible variations.

In 1950s, William Deming's studies regarding quality, statistical quality control techniques inspired by Shewart, have been used in the military known as the American War Standards. In following years, Deming proposed some quality improvement standards, concept of variations to the Japanese. He developed a systematic approach to problem solving known as the Deming or PDCA (plan, do, check, act) cycle. His quality studies proved that quality is also increasing productivity. Deming is considered as the originator of the modern quality movement.

After 1950s, some statistical tools also designed by Joseph M. Juran in order to improve quality and to assure the products have standards in point of customers. Juran is considered as "the most important contributor to quality management after Deming when publishing his book "Quality control handbook" in 1951. Juran highlighted managerial responsibility for quality as Deming pointed (Westcott, 2013). He claimed that quality control must be applied as an integral part of management. In addition, he claimed that customer needs have to be taken into account, and defined quality as 'fitness for use'.

Another contributor of quality is Armand V. Feigenbaum, the originator of Total Quality Control. He developed a systematic approach for quality improvement and defined the quality as 'best for the customer use and selling price'. He defined the quality from a customer's perspective the term Total Quality Control originated from his "Quality control: principles, practice and administration" book. Feigenbaum claimed that quality should be taken into account at an early stage, rather than inspecting and controlling during production.

Kaoru Ishikawa used concept of Total Quality Control with concepts of Deming and Juran and translated, integrated and expanded these concepts into the Japanese system. His approach regarding quality is starting from top management to lower-ranking employees so that company could be successful and sustainable.

Taguchi, Japanese quality expert, developed an approach to optimize quality at the design stage. He developed the concept of the Quality Loss Function which focused on quality loss rather than quality. He explained the quality loss as "loss imparted by the product to society from the time the product is shipped". His Quality Loss Function showed a reduction in variability, increasing in quality. His estimation was that 80% of all defectiveness caused by poor design, and he took the concepts of quality back to the

design stage so that the noise variables which disrupt production can be eliminated (Westcott, 2013).

Philip B. Crosby developed an approach called ‘zero defects’. Even though, ‘zero defects’ was an approach improving quality, that does not mean workers never make mistakes, but also the company does not expect them to make mistake during their process. He suggested the same idea with Deming; 85% of quality problems can be stopped by management control. In 1979 ‘Quality is free’ in 1984 ‘Quality without tears’ has been published by him.

Another contributor to the quality concept was Garvin who believed that if quality can be managed, first it must be understood. He published “Managing quality” (1988) and many more articles. Garvin have identified and examined the quality as eight critical dimensions. Garvin was one of the first people, who focused quality with its critical dimensions which is widely adopted throughout the world.

Today, TQM is still an important quality improvement standard for any organization. With the introduction of many quality improvements approaches such as Six Sigma, and ISO 9001, TQM makes any organization or process of any production much better.

2.2.2 Service Quality

After 1950s, quality discussions have been changed from product quality to service quality due to the operations in service industry has become more important than product manufacturing.

In 1960s, discussion was going around definition of service and differences between services and products. One of the first people who defined differences among service and product was Regan who claimed that “intangibility, perishability, heterogeneity and ubiquity make the total comprehension of services difficult” (Regan 1963, p. 58).

The primary research was trying to determine what service quality meant to customers (Zeithaml al., 1996). Zeithaml claimed that due to characteristic of the service, it was hard to define and evaluate it (Zeithaml, 1981). Since in the previous research quality has been defined multi dimensional (Parasuraman, 1988) there was no agreement about how to evaluate the service quality (Cornin and Taylor, 1992).

Because of lack of consensus, service quality is most debated subject in service industry (Gupta and Chen, 1995).

There are many different service quality models has been developed in various industries i.e. retailing and servicing industry. In 1982, Grönroos also defined service quality dimensions as functional aspect and technical aspect. Later ServQual by Parasuraman has been developed (1988). Later, WebServQual, an extension of ServQual, were proposed (Li et al., 2002).

Among all models, ServQual which was defining and measuring service quality was the most cited and discussed article. ServQual had a significant impact on quality of service in literature and industry as well. ServQual was measuring performance (P), customer expectations (E) and quality as follows Q=P-E.

After these studies, in 1997 Berry's and Parasuraman published another article regarding quality of system "Listening to the Customer-The Concept of a Service-Quality Information System," This was another remarkable study which encouraged organizations to measure the quality of their customer service. An effective service-quality information system (SQIS) has to periodically survey not only their customers, but also employees and competitors' customers.

2.2.2.1 Measuring Service Quality

As stated before, even though service quality was hard to be defined and measured, various researchers defined service quality and tried to measure it (Lewis and Booms 1983Grönroos 1984, Parasuraman et al. 1985 and 1988, Carman 1990, Cronin and Taylor 1992, Teas 1993, Westbrook and Peterson 1998).

Lewis and Booms (1983) looked at the service quality in customer perspective and claimed that service level delivered to customer has to be matched with customer expectations. Grönroos (1984) claimed that consumer's measure (perceived) service quality by comparing their expectations with experiences of the service that they have received. In addition, Parasuraman(1988) pointed out that "perceived service quality is viewed as the level of discrepancy between consumers' perceptions and expectations".

2.2.2.2 Service Quality Measurement Models

Grönroos (1984) developed a model and e expressed that consumers were comparing the service experienced with the service expected when measuring service quality. The model attempted to understand how the quality of service given perceived by customers (quality from customer perspective).

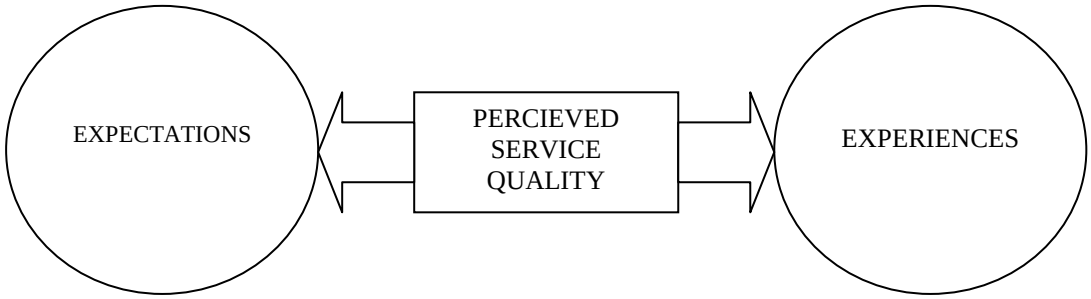


Figure 2.2 Grönroos's perceived Service Quality model.

If customers are happy about the service experienced, expectation of customer is fulfilled. In other words, the experienced quality meets the expectations of the customer; which is called the expected quality.

Grönroos's model was very important and widely used because it still reminds any company that service quality is not just part of production or service but also the way how it is delivered to customers in crucial as well.

One of the most discussed service quality measurement model can be seen in Figure 2.3 as conceptual model of service quality. There are 5 gaps between each level in the model. These gaps represent a group of problems which need to be fixed by service providers in order to satisfy customers.

These gaps can be fixed easily if providers consider the problems. For instance, first gap is regarding management perceptions of consumers' expectation and customers' expected service. The reason behind this gap could be lack of quality research which should be done by university. If service provider searches for what customers want, they might fulfill the customers' expectations. The other gaps can be seen be seen in below.

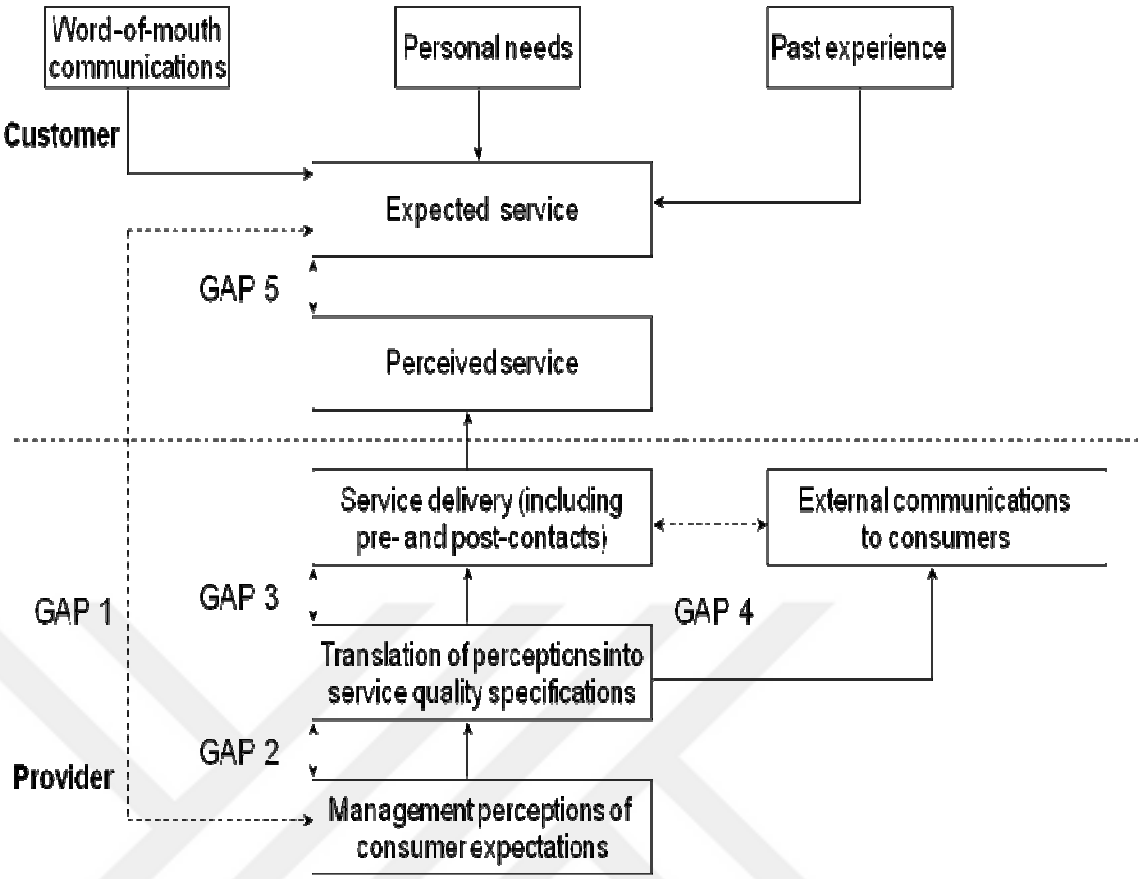


Figure 2.3 GAP model of service quality (Parasuraman et al. 1985).

Although, ServQual has been improved by Zeithaml et al. (1990), used widely by many researchers and it has been criticized by many authors as well in terms of reliability and validity of ServQual scale (Cronin and Taylor 1992, Teas 1993).

2.2.3 A Brief Overview for Human Computer Interaction and Usability

Usability firstly has been defined under the human–computer interaction (HCI) studies. Usability has been accepted key factor designing the user interface (Shneiderman, 1998). The core usability concept and studies started with the discussion of HCI after 1980s.

Human Computer Interaction has been started at 1960s when computers were scarce, expensive, bulky machines used for automatic calculations. Douglas Engelbart saw that computer has a potential to be as personal interactive tools and he started to improve computing hardware and software in order to increase the human intellect;

capabilities of users. Even though their perspective was inspiring, complex of the computer was hard to understand and their ideas somehow could not exist.

Later at the Xerox Palo Alto Research Center (PARC) in the 1970s, Engelbart and other researchers developed the hardware and software, which was less complex for novice users. They influenced many developments such as mouse-based interactions, direct manipulation interfaces.

The contributions to user interface by Apple Macintosh and Microsoft Windows in 1980s, lead some development on interfaces. First HCI was regarding graphic design and basic programming but later some practitioners realized that interactive software design is more important for them. Interactive system was supporting multimedia, multitasking which is more important than graphical design. After realizing this, researchers changed their studies to intelligent user interfaces Many HCI studies have been changed “from computers and graphical interfaces to usability evaluation, interaction design and user experience” (Roussel, 2014).

HCI practitioners were interested how users interacted with the computer, and "...long lists of guidelines of good practices" were introduced (Dumas, 2007, p. 55). At the end of 1980s, the "usability engineering" term has been used in the literature and practitioners conducted user testing and collected quantitative data from users (Dumas, 2007, p. 55).

In 1980, Ericson and Simon published " Verbal Reports as Data" which focused on using the Think Aloud Method has been introduced and this method used widely in usability evaluation of product or services.

In1984, Macintosh computer interface has been introduced by Apple which has been influenced by Xerox technology. In addition, “997 guidelines for Designing User Interface Software” was published by Smith and Moiser. Harry Hersh and Dick Rubinstein wrote “The Human Factor” the first book-length description of human-computer interaction.

In 1987, Ben Shneiderman published the “Designing the User Interface” book. Also, he published “The Questionnaire for User Interaction Satisfaction (QUIS)” based on the work at HCI lab at the University of Maryland. The method has been cited and used in usability studies.

In 1988, some papers, books have been published regarding usability (Whiteside, Bennett, & Holtzblatt, 1988). These publications were pointing out early goal setting, prototyping and iterative evaluation.

1990 is a mile stone for usability because root of today's publication has been introduced in this year. "Human Factors and Usability" which defined usability as a function of efficiency, effectiveness & satisfaction (the ISO 9241 pt 11 standard) has been published by Shackel. He made enormous contribution to usability methods because during last 3 decades many definitions have been introduced regarding usability however usability is still considered mainly in terms of these three aspects. Additionally, "Heuristic Evaluation of User Interfaces," has been published by Jakob Nielsen Rolf Molich. This article brought many contributions to usability evaluation method and accepted one of the most influential usability methods.

In 1996, System Usability Scale (SUS), one of the most used and cited method, was published by John Brooke.

In 1998, usability became a standard in ISO 9241. One of a new wave of books using the term User Experience, "Web Navigation: Designing the User Experience" was published. "The Evaluator Effect in Usability Tests: Problem Detection & Severity Judgments--the first study document the "evaluator effect has been published by Jacobsen, Hertzum, and John.

In 2000, Steve Krug published "Don't Make me Think". Method brought usability testing to the masses using the same Think Aloud method from Ericson and Simon from 20 years earlier.

In 2002, "An Empirical Comparison of Lab and Remote Usability Testing of Web Sites have been published by Tom Tullis. It was the one of first publications about remote usability testing.

In 2010, Beyond the Usability Lab: Conducting Large-Scale User Experience Studies has been published.

In 2012, UPA was changed to the User Experience Professionals Association (UXPA).

2.2.4.4 Relation between Usability and Quality

Some researcher defined usability as quality in use, ultimate quality factor. There are many studies found that usability and quality is related with each other and affects each other. (Bevan, 1995, 1999; Folmer and Bosch, 2004; Seffah et al., 2008) Since usability is accepted as quality according to some researchers, they claimed that usability could be applied to any products or services such as to traffic signs, cameras, books, alarm clocks, computers, usually to software applications, development and websites (Rubin and Chisnell, 2008, Krug, 2006, Abran et al., 2003).

Therefore, quality of service become very important factor and has been found to be one of the most important factors for online transactions (Kim and Lee, 2002). In another study, Liang and Lai found that design quality of interface significantly affected consumers' choices (Liang and Lai, 2000). In a recent study, Ladhari defines usability as a component of quality for services (Ladhari, 2010) and Mack and Sharples defines usability as a critical dimension of product quality which affects product success (Mack and Sharples, 2009).

As seen above, methods to measure the quality and usability have been discussed for several decades, and interestingly, there has been perception that usability and quality is related with each other and sometimes, they have been used instead of each other (Oztekin et al., 2009). It is known that "usability" appeared after product improvement process and quality studies.

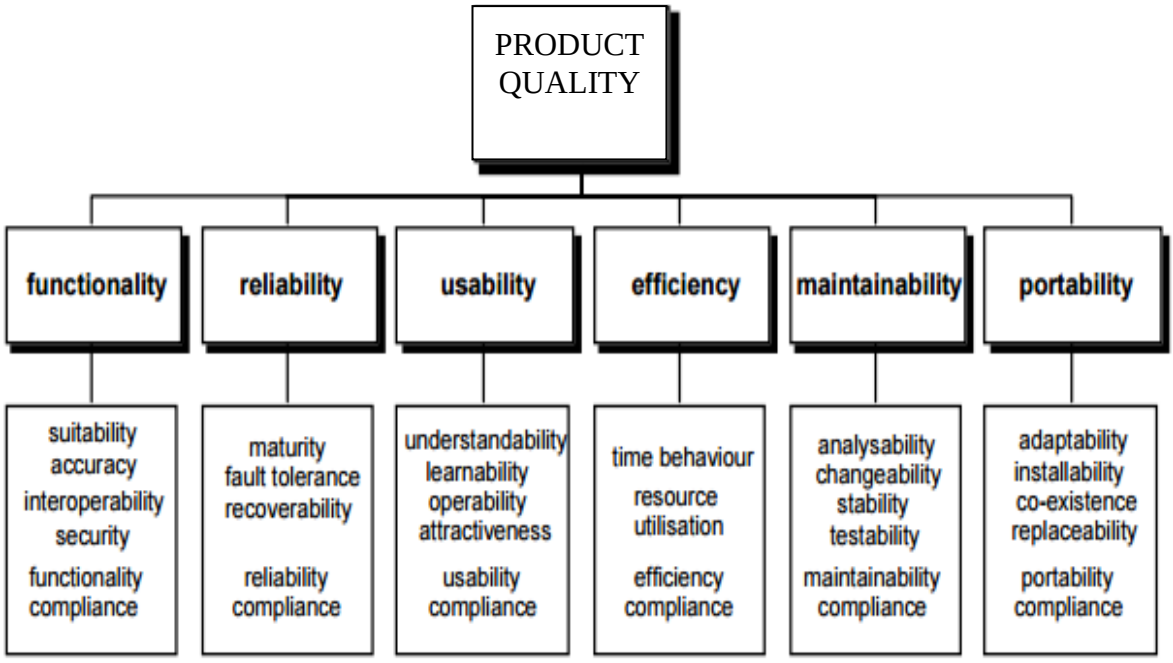


Figure 2.4 ISO/IEC Product quality model.

The ISO/IEC 9126 determines usability as one of six characteristics (factors), under product quality model. Additionally, ISO/IEC 25010:2011 defines usability as sub-characteristic under quality.

2.2.5 Literature Review for Usability

Usability has a long history adventure before known all around world with it's today popularity. It has been first mentioned in the Palm Beach Post in 1936, later during some product development process from 1943 to 1956 known in quality improvement process, and after 1959, under ergonomics science, also known as quality of computer systems (Shackel, 1959) and ease-of-use (Miller, 1971; Bennett, 1972), and later in 1979, usability has been used with human computer interaction heading together.

"The Commercial Impact of Usability in Interactive Systems" has been written by John Bennett under the heading usability. In 1988 Ericson and Simon published "Verbal Reports as Data" which focused on using the Think Aloud Method that is used widely in usability evaluation of product or service.

In 1985, "Designing for Usability: Key Principles and What Designers Think" has been published by J. Gould and Clayton Lewis. It was the first time user interface design has been discussed. In 1988, Ben Shneiderman published the "Designing the User Interface" book.

He also published "The Questionnaire for User Interaction Satisfaction (QUIS)" based on the work at HCI lab at the University of Maryland. After these studies, 1988 is history of usability life cycle birth stage for the point of modern usability profession. Usability is a core term in human–computer interaction (HCI), (Hornbaek 2006). Since human and computers (machines) interacted with each other, usability is part of human computer interaction.

Number of studies related to web usability from 1997 tremendously has increased year by year and number of researches on in the field of usability evaluation including web based usability evaluation were more than relative areas such as human computer interaction, software and web engineering by researches (Fernandez et al., 2011). While

usability studies have been increasing, definition of usability has been changed in its history, and there are many different definition of usability in literature currently.

Usability is defined as ‘‘the capability to be used by humans easily and effectively’’ (Shackel, 1991, p. 24); "...the ease of use and acceptability of a system or product for a particular class of users carrying out specific tasks in a specific environment"(Bevan, 1991), “designing software applications which people find convenient and practicable for use “(Nielsen, 1993) ‘‘quality of use’’, and “quality in use” (Bevan, 1995, 1997). Keevil described usability in computer science as “how easy it is to find, understand and use the information displayed on a web-based system” (Keevil, 1998).

Usability means that people who use a product can do so quickly and easily to accomplish their own tasks. This definition rests on four points: (1) Usability means focusing on users; (2) people use products to be productive; (3) users are busy people trying to accomplish tasks; and (4) users decide when a product is easy to use (Dumas and Redish, 1999). Usually definition of usability is regarding context dependent (Newman and Taylor, 1999).

Other definitions of usability with respect to years as follows:

How well and how easily a user, without formal training, can interact with an information system or web site [where the information system or website is tested and not the user]” (Benbunan and Fich, 2001);

"How well the intended users can interact with technology to carry out an assigned activity", "an interface that is workable and intuitive from the user's point of view", "usability research strives to improve both the efficiency and effectiveness of systems" (Hallahan, K. 2001);

“Ease of use” something that can be done easily the way how it is indented to and can be applied to any object which is used for some purpose (McNamara, 2003);

“The effectiveness, efficiency, and satisfaction with which specified users can achieve goals in particular environments” (Hornbaek, 2006);

"Usability really just means making sure that something works well: that a person of average (or even below average) ability and experience can use the thing—whether it's a Web site, a fighter jet, or a revolving door—for its intended purpose without getting hopelessly frustrated" (Krug (2006);

Strategic factors for software development (Juristo et al., 2007);

“The ultimate quality factor” for software architecture (Seffah et al., 2008); "the degree to which a given piece of software assists the person sitting at the keyboard to accomplish a task, as opposed to becoming an impediment "(Levi and Conrad, 1998);

The level of easiness and difficultness for systems which will determine the users’ success or failure (Insfran and Fernandez, 2008).

Usable systems provides many advantages such as increasing productivity, reducing error (Lallemand, 2011, p. 299) and increasing comprehension of the content (Flavian et al., 2006) and satisfaction for the users. Poor usability systems can cause more time to learn complex systems and more effort to correct the mistakes.

Usability could be applied to any products or services such as to traffic signs, cameras, books, alarm clocks, computers, usually to software applications, development and websites (Rubin and Chisnell, 2008, Krug, 2006, Abran et al., 2003).

Nielsen recently defined usability, as “Usability is a quality attribute that assesses how easy user interfaces are to use”. Also he claims that "usability" stands for to improving ease-of-use during the design process (Nielsen, 2012). According to Nielsen usability has 5 principles which define usability, Learnability, Efficiency, Memorability, Errors and Satisfaction.

Learnability: when first time they visit the web site, how easy to accomplish basic tasks?

Efficiency: How quickly do users do things on web site once they learned the design?

Memorability: Is it easy to remember things on web site in future visit?

Errors: How many errors do users make and how severe they are? In addition, how easily can they make themselves correct?

Satisfaction: Are users satisfied enough after using the web site?

However, Whitehead (2006) provides different principles for the usability of a website:

Checkability: The system should allow the users to make sure the correct information is going in and going out of it.

Confidence: Users should have confidence while using the system in terms of their capability.

Control: Users have control over the system, especially when entering information and when taking information out of the system.

Ease of Use: Easiness of the system while using.

Speed: The system should be fast enough to be used.

Understanding: The system should be clear and content understandable.

Definition of the International Standardization of Organization, (ISO) is accepted one of the most common definition of usability in literature, which is ‘‘the extent to which intended users of a product achieve specified goals in an effective, efficient and satisfactory manner within a specified context of use’’ (ISO 9241). According to ISO (international standardization of organization), it is not possible to measure usability directly because usability must be considered as it has three components in order to be measured. Other ISO definitions are as follows:

Table 2.1 Definition of ISO for usability.

Standard	Definition
ISO 9126-1 (2000)	Under defined conditions, the capability of the software product to be understood, learned, and used attractively.
ISO 9241-11 (1998)	“The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use.”

IEEE 600.12 (1990)	Ease of systems which allows user to learn, operate and prepare the systems or components.

In addition, usability defined as “it is the property of the system that defines its degree of simplicity of use in terms of learning, storage and efficiency” (Cassino et al., 2014), and “usability is the performance achieved and satisfaction experienced by system users” (Wagner et al., 2014)

After all, usability can be defined as quality for the outcome of human and computer interaction. In addition, usability aims to remove obstacles between users and their desire; so that users achieve their purposes quickly, complete their tasks correctly as expected; and get satisfied while interacting with machines, products, or services. In addition, usability term has been used usually with software development, especially in web-based information systems. Furthermore, usability term has been used usually with software development and especially in web-based information systems. There is no a final definition for usability term because usability reflects a continue process of improvement over the time. Today is not usable for tomorrow because tomorrow requires more than before. For example, first computers and cell phones; they have been improved dramatically over the time from button-based to touch-based screen. Today button-based devices are not being used widely because people preference changed dramatically towards to more usable systems.

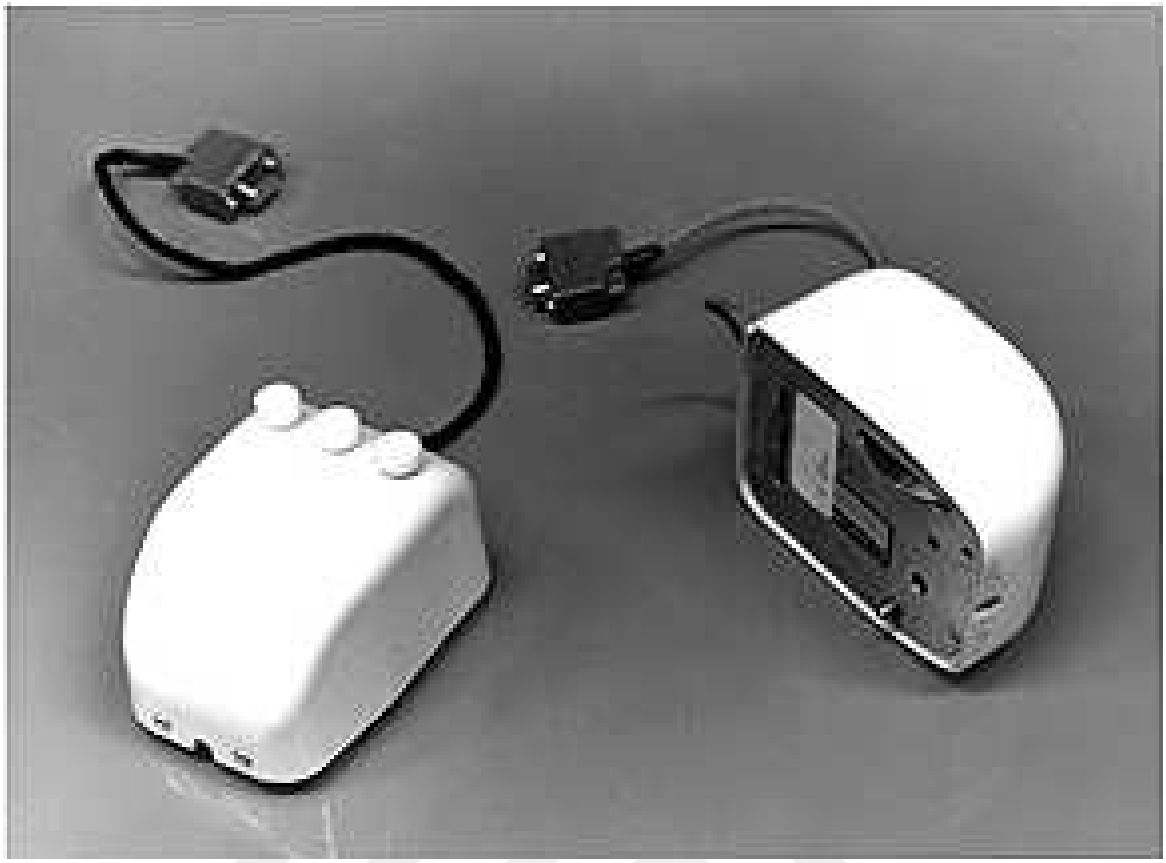


Figure 2.5 An example of a mouse used in 1968 (Andrew, 2016).

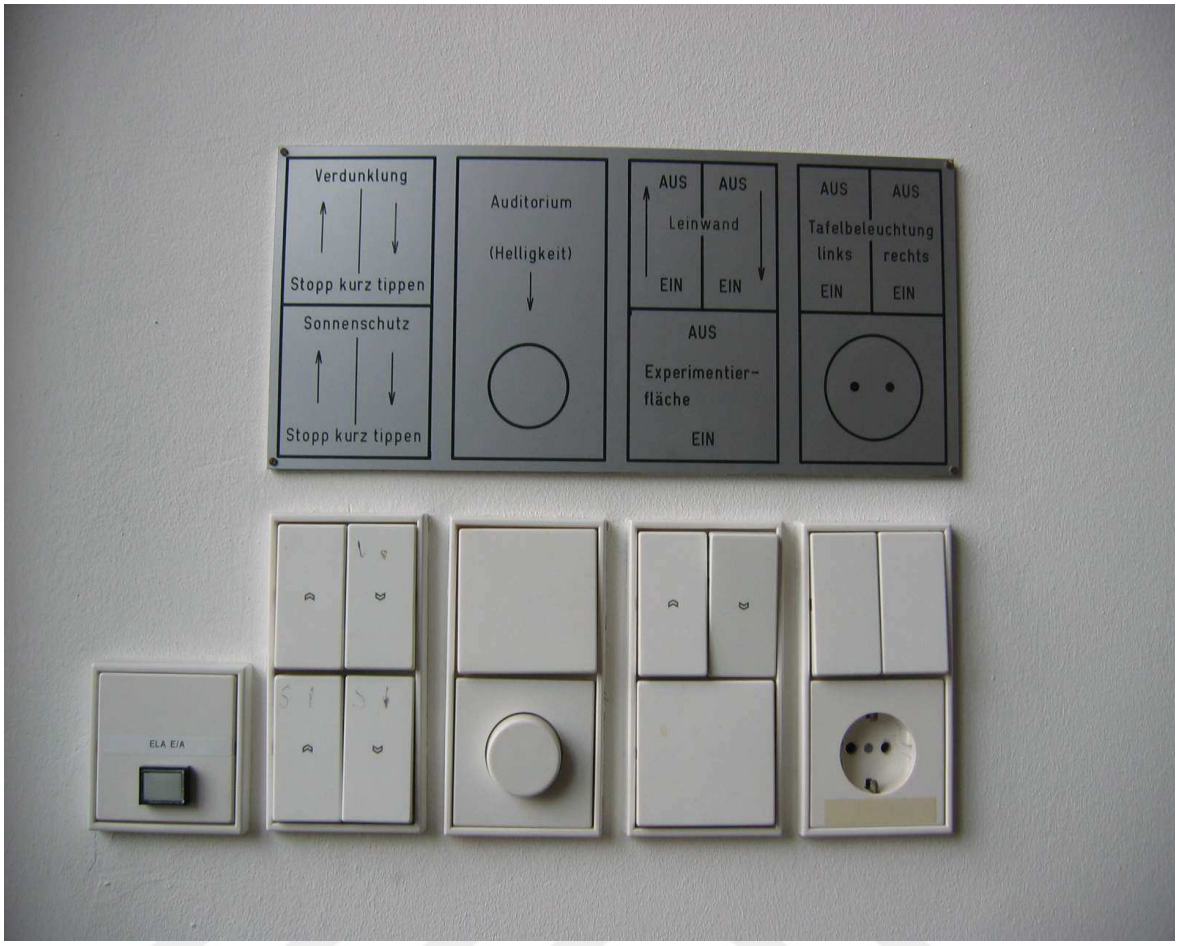


Figure 2.6 An example of a control panel. (Andrew, 2016).

Over the years products such as control panels, mouse pads, and computers have been changed dramatically.

Table 2.2 Definition of usability over the years.

1. ‘‘the capability to be used by humans easily and effectively’’ (Shackel, 1991, p. 24)

2. "...the ease of use and acceptability of a system or product for a particular class of users carrying out specific tasks in a specific environment "(Bevan, 1991)
3. “designing software applications which people find convenient and practicable for use” (Nielsen, 1993)
4. It is important to realize that usability is not a single, one-dimensional property of a user interface. Usability has multiple components and is traditionally associated with these five usability attributes: learnability, efficiency, memorability, errors, satisfaction. – (Nielsen, 1993)
5. “how easy it is to find, understand and use the information displayed on a web-based system” (Keevil, 1998)
6. "the degree to which a given piece of software assists the person sitting at the keyboard to accomplish a task, as opposed to becoming an impediment "(Levi and Conrad, 1998).
7. Usability means that the people who use the product can do so quickly and easily to accomplish their own tasks. This definition rests on four points: (1) Usability means focusing on users; (2) people use products to be productive; (3) users are busy people trying to accomplish tasks; and (4) users decide when a product is easy to use. - (Dumas and Redish, 1999)
8. “a measure related to how usable or user-friendly the product, service, or system is “(Flowers, 2000)
9. “how well and how easily a user, without formal training, can interact with an information system or website [where the information system or website is tested and not the user]” (Benbunan and Fich, 2001, p. 151)
10. how well the intended users can interact with technology to carry out an assigned activity", "an interface that is workable and intuitive from the user's point of view", "usability research strives to improve both the efficiency and effectiveness of systems" (Hallahan, K. 2001)
11. “ease of use” something that can be done easily the way how it is indented to and can be applied to any object which is used for some purpose (McNamara, 2003), “the effectiveness, efficiency, and satisfaction with which specified users can achieve goals in particular environments” (Hornbaek, 2006)
12. "usability really just means making sure that something works well: that a person of average (or even below average) ability and experience can use the thing—whether it's a Web site, a fighter jet, or a revolving door—for its intended purpose without getting hopelessly frustrated" (Krug , 2006)
13. “the list of strategic factors to be dealt with, especially in software development” (Juristo et al., 2007)
14. “the ultimate quality factor” for software architecture (Seffah, 2008)
15. "the ease or difficulty that user's experience with systems, [such as web applications], will determine their success or failure" (Insfran & Fernandez, 2008, p. 81)
16. Nielsen recently defined as “Usability is a quality attribute that assesses how easy user interfaces are to use”. Also he claims that "usability" stands for to improving ease-of-use during the design process (Nielsen, 2012)
17. ‘‘the extent to which intended users of a product achieve specified goals in an effective, efficient and satisfactory manner within a specified context of use’’ (ISO 9241)
18. “It is the property of the system that defines its degree of simplicity of use in terms of learning, storage and efficiency” Empirical validation of an automatic usability evaluation method. (Cassino et al., 2014)

2.2.5.1 Usage of Usability Techniques

Usability techniques are used widely in different research areas such as in health industry, web sites development, software improvement, products, mobile phones, automotive industry, and so on.

Here are some research areas where usability techniques have been used recently:

In health industry for evaluating the usability of application or devices (Kaufman et al, 2006), mobile health technology (Brown et al, 2013, Leuthold et al., 2007), designing web site for blind people (Leuthold et al., 2007), user buying preference for shopping (Lee and Kozar, 2011), web mining to find out user preference (Wu et al., 2014), Vehicle Information Systems (Harvey et al, 2010), smartphone interface (Choin and Lee, 2011), electronic shopping (Chen and Macredie, 2005), learning for disabled students (Yussof et al., 2013), e-Learning sytem(Masood and Musman, 2015, Biasutti, 2011), web sites usage for elderly people (Castillaa, 2015), software developmet with designer of the web site (Bruun and Stage, 2015), impact of the age on web usage (Wagner et al., 2014), governmental web site usability (Elling et al., 2012), university web sites (Sengel, 2013, Turan and Bayram, 2013).

2.3 WEB SITE BASICS FOR USABILITY

It is not easy to define the basic features of a usable web site because it depends on many factors. However, the most important way to find out what a usable web site is that whether it has customer's perspective in designing of the web site, namely users has been considered when the web site designed or not. In order to put the customers in designing of the web sites mean that service providers have to know the customers, who they are, what they expect from the web site, their preference. A user preference is choice of alternatives, which alternative is better than other and it shows the feeling of users. There are many factors that make web sites usable for its users; such as quality, price, brand, performance and so on (Lee and Koubek, 2010).

According to Nielsen, usability has 5 principles which define usability; Learnability, Efficiency, Memorability, Errors and Satisfaction.

Learnability: when first time they visit the web site, how easy to accomplish basic tasks?

Efficiency: How quickly do users do things on web site once they learned the design?

Memorability: Is it easy to remember things on web site in future visit?

Errors: How many errors do users make and how severe they are? And also how easily can they make themselves correct?

Satisfaction: Are users satisfied enough after using the web site?

Flavia´n 2006 says that design should be with respect to (a) the ease of structure though website (b) simplicity of use though website (c) the speed which the users can find as quickly as possible what they looking for; (d) the perceived ease of site navigation, less time to accomplish desired, (e) the user control over the web site.

Time to learn to use the site should as less as possible, close to zero; otherwise the site could disappear (Nielsen, 2000).

Web sites should provide help and never use homonyms, synonyms, misspelled and confusing words. Wording has to be selected with respect to its purpose not for only designer perceptive but also to the actual target users.

Documentation is supposed to be designed logically easily reachable. If the navigation makes the web site difficult to use, users/customers will go away.

The links and pages must connect with respect to desired purpose; wrong link will make users not satisfied. Even if web site has the right information in somewhere, if the user does not know where it is or how to find it, this could user unhappy and users will leave the web site (Marsico and Levialdi, 2003).

Also some researchers offered different factors for a usable web site such as content, format, search-capability, aesthetics, and speed of web site (Lee and Kozar, 2012).

2.3.1 Web Site Usability

People mostly do their daily activities thru the web sites such as money transferring, online shopping, and watching their favorite shows and so on. However, developing a usable web site for customers is one of the most complicated issues in online world. If customers are satisfied with service experienced thru the web sites, they visit the web sites again. Dissatisfaction of the web site, due to poor interface, could result an unacceptable learning time of the content, too many clicks to navigate, and longer time of the web pages loading. If users have been offered more usable web sites by competitors, they might visit those web sites (Ratner, 2003). According to Nielsen, if customers cannot find what they look for, then they will not buy it (Nielsen, 2000).

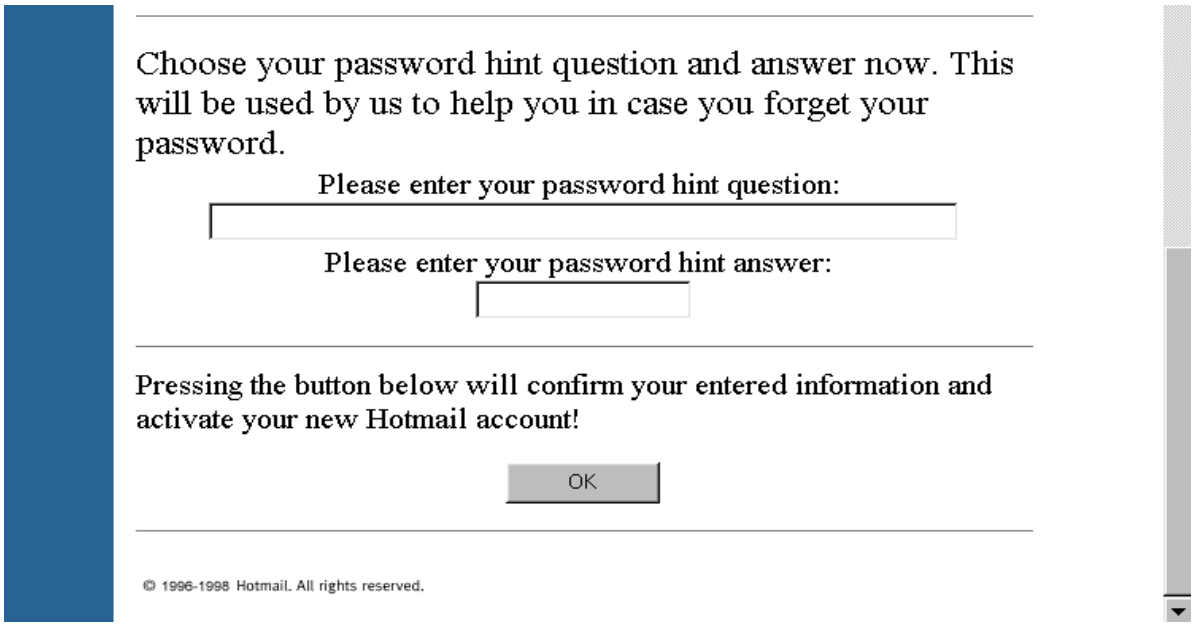


Figure 2.7 Part of the Hotmail registration process before usability testing (Andrew, 2016).

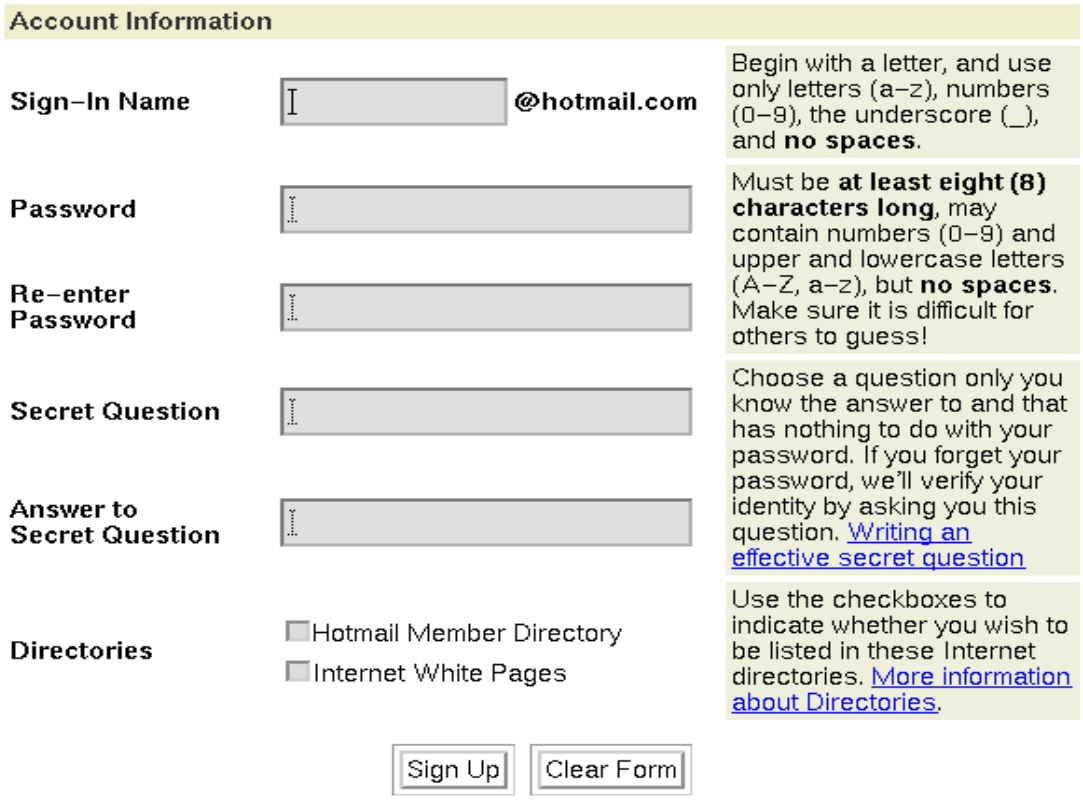


Figure 2.8 Design of the Hotmail registration after usability testing (Andrew, 2016).

The dialogue was redesigned as a “secret question” which is more logical word. Additionally, registration process explained clearly for each step.

A research has found that if customers cannot find what they are looking for, potential sales will approximately drop by %50 or more. Additionally, if customers had a negative experience when they visit web site at first time, %42 of them do not want to return the web site again (Manning, 1998). In addition, Becker and Mottay expresses that users are running away from unusable sites which causes online business failures.

Usability of web sites is sometimes regarding customer perception for the service experienced. Ben Sheiderman found that if web site change designs of interface such a color, position of elements, terminology of button, this could decrease the performance of the web site up to 25%.

Majority of the web site sites violated the simple design principles. There are some evidences regarding how usability affects users’ behaviors as follows:

15 commercial web sites analyzed; even though users the test from correct home page, only 42% of the information was found by users.

62% of web shoppers gave up looking for the item they tried to purchase online.

51% compliance were with simple web usability principles such as "is the site organized by user goals?" and "does a search list retrievals in order of relevance?"

What do they do when they could not find what they are looking for or what if they are not satisfied with the service they took from a web site? They would rethink whether or not it's right place, “Google it”, or visit the competitors web site.

Before the usability concept, designer developing the web sites without thinking if there could be a problem for users. Spool (2009) claimed that after usability testing, there are many usability problems have been found and fixed which increased satisfaction of users and sales. The issues sometimes were just only removing some disturbing process for registration or changing the whole design of the web site but only adding search button or improving some section of web sites. For example, IBM reports that after redesigning of the web site, usage of search engine in the web site has been decreased %84 by users because the most used feature of web site was search engine. After redesigning the web site, the sales were increased by 400% because users were finding easily what they were looking for without using the search tool as they did before (Tedeschi, 1999).

Users want the web sites to be more usable, easy to use and understand the content quickly. In addition, they are interested some special features such as download speed, trust, responsiveness, and empathy (Downing and Liu, 2011)

Developers, researchers, and companies have seen the importance of usability studies. Therefore, the number of studies regarding web site usability has been increased since 1997. As it can be seen in the Figure 2.3, number of usability studies has been increasing over the years.

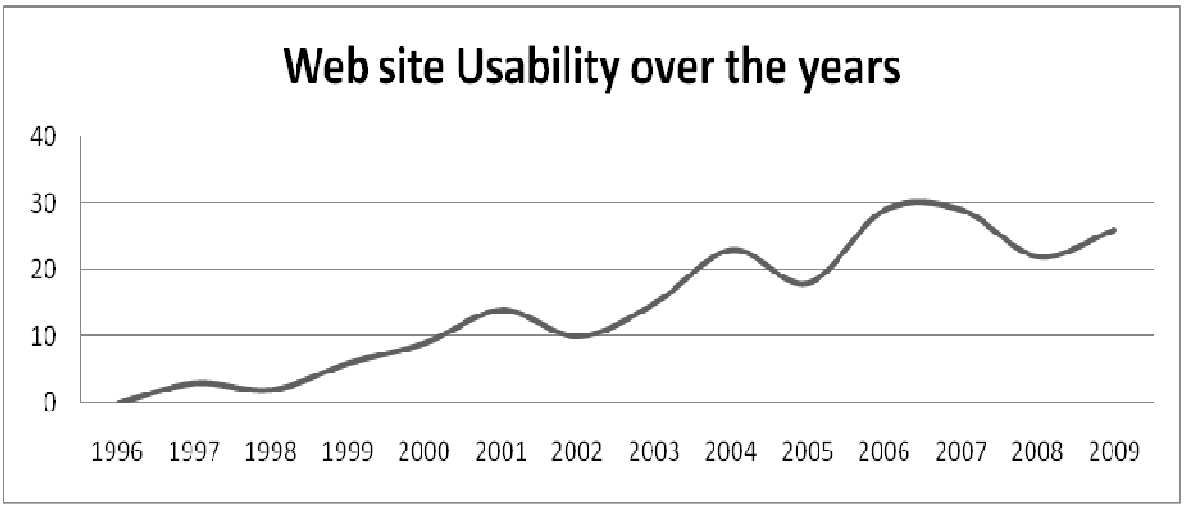


Figure 2.9 Number of articles regarding web site usability over the years (adopted from Fernandez et. al., 2011).

The importance of usability studies resulted new techniques and methods for usability measurement. In the next section, some usability evaluation techniques will be reviewed.

2.4 USABILITY EVALUATION TECHNIQUES

There are many techniques used to improve the usability of interactive systems. Some of these methods measured the usability of web sites using guidelines and heuristics while others used different techniques such as questionnaires and user testing. These techniques can be classified as follows: guidelines for improving usability of a system (Smith and Mosier, 1986), heuristics to evaluate user interfaces (Molich and Nielsen, 1990), methods to predict usability problems (Walji et al., 2014), discussions on how to measure usability (Nielsen and Levy, 1994; ISO, 1998; Hornbaek, 2006, Walji et al., 2014; Fogli and Guida, 2014) and checklists to evaluate usability (Oztekin, 2009; Lee and Kozar 2011, Elling et al., 2012). Some usability techniques which are used common in literature are shown below in Figure 2.2. Some usability methods used in literature will be explained briefly.

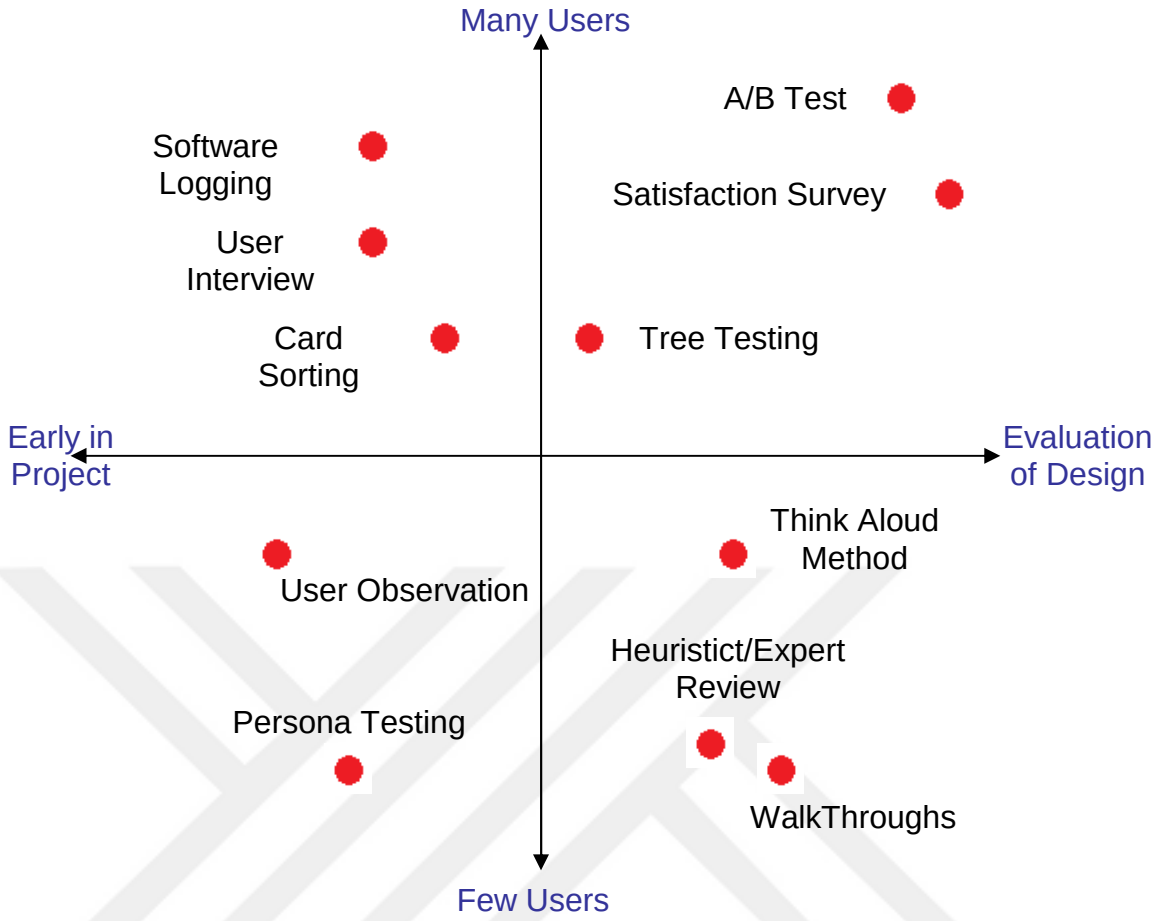


Figure 2.10 Usability techniques and their usage (Adopted from Quesenbery, 2008).

As seen in the Figure 2.2while some usability techniques can be applied with a few users, some needs more users. In addition, techniques can be applied early in project; some of them can be applied evaluation of design.

Moreover, in Table 2.2, the methods can be divided in terms of techniques (inspection or test), type of data (quantitative or qualitative), the place where they are in life cycle, (beginning, design, after finish, or anytime-during process) and so on (Schriver, 1989). Furthermore, Table 2.2 provides more details for each method. In other words, the number of users needed, stage of life cycle, and type of data to be collected for each method can be seen in Table 2.2 in detailed format. For instance, a survey method can produce both qualitative and quantitative results. In addition, surveys can be used anytime in stage of life cycle. Detailed information for each method can be seen in the table.

Table 2.3 Usability methods with their features (Adopted from Andrews, 2016).

METHODS	Stage of Life Cycle	Qualitative/ Quantitative	Inspection / Test	Number of Specialists/ Experts	Number of Novice Users	Explanation
Software Logging	Design	Quantitative	Test	1 to 2	20+	Better in early of project to collect data
User Interview	Design	Qualitative	Test	1 to 2	>3	Results are very satisfactorily
Card Sorting	Anytime	Quantitative	Test	3 to 5	100s	Can be applied to anyone either expert or novice
Tree Testing	Anytime	Quantitative	Test	3 to 5	100s	Better after card sorting to check new design
Persona	Beginning	Quantitative	Test	2 to 12		Generally 7-8 people are enough
User Testing/ Experimental	Finished System	Both	Test	2 to 3	50+	Results are very satisfactorily, and *usually* Quantitative
Remote Testing/ Experimental	Finished System	Both	Test	2 to 3	50+	Results are very satisfactorily, and *usually* Quantitative
Heuristics Evaluation	Anytime	Qualitative	Inspection	3 to 5	NA	Results are satisfactorily
Walkthroughs	Anytime	Qualitative	Inspection	3 to 5	NA	Results are satisfactorily
Think Aloud Method	Design	Qualitative	Test	2 to 4	3 to 5	Results are satisfactorily
Surveys	Anytime	Both	Test	3 to 5	100+	Results are satisfactorily
A/B Testing	Finished System	Quantitative	Test	2 to 3	1000s	Results are satisfactorily
First Click Testing	Anytime	Quantitative	Test	10s	100s	More people leads powerful data

Researchers have used a different terms in order to describe usability testing methods. According to Nielsen and Molich (1990), there are four terms to measure usability: Formally, Automatically, Empirically, and heuristically. Redish et al. (2002) defined usability testing as Automated, Inspection (a heuristic evaluation or cognitive walk-through), Performance (where users complete tasks), and Operational (system information such as download speed).

In addition, Whitehead (2006) determined usability testing as Inquiry (surveys, focus groups, interviews etc.), Inspection (heuristic evaluation or cognitive walk-through), and Formal (list of task to complete). According to Hallahan (2001), usability-testing methods are laboratory testing and alternative assessment (outside the laboratory setting).

Traditional usability tests have been conducted in a laboratory environment by an evaluator. Usability testing can be applied to developers, human-computer interface experts, or representative end users (Levi and Conrad, 2008). Laboratory testing is a classic and one of the most reliable techniques in order to measure usability.

Krug's (2006) define usability testing in laboratory simply as follow:

If designer of a web site or owner of a product wants to know, whether the web site or software is easy to use, simply they need to watch some people and take notes while they use it. If the observer saw that users face with issues while using it, then designer or owner should fix it.

When usability tests have been used in 1990s, it was a very expensive method. First, a usability lab with an observation room was necessary. In addition, it was believed that at least two video cameras in order to record the users' reactions while took the tests and hiring many people were necessary which could cost from $20,000 to $50,000.

When websites are evaluated in laboratory environment, screen capture software is used to record mouse movements (Hallahan, 2001) and with these special programs, users can be evaluated in terms of effectiveness and efficiency as well as mention above. For example, whether they were able to complete the tasks shows effectiveness while how quickly the tasks were completed shows efficiency.

He also explains that in the laboratory evaluator takes scoring sheets while watching the users. After data collected, it will be combined with the camera recordings, comments from the participant, questionnaires, and mouse-click data. In addition, a timer can be used while users start the task in order to see how quickly they completed the task assigned. "Together these multiple measures provide richer insights into the user's response than would be possible with a single measure" (Hallahan, 2001, p. 226).

Below, there is a summary of the usability procedures in a laboratory environment,

1. Identifying the problem, developing the test question, explaining the purpose of study

2. Determining what is going to be measured and which method is going to be used,

3. Conducting a pilot test with evaluators and a couple users identify what participants and evaluator will do during the test,

4. Selecting and planning measurements, defining tasks, writing the scenario, deciding the equipments, sheets, camera recording devices, questionnaire, and special tool for screen movement during the test, and number of evaluators,

5. Identifying the population and inviting them to the test

6. Collecting data during the test based on observation, recording devices, mouse movements, screen movements, timer, face impression, questionnaire, and at the end of the test feedbacks from users.

7. Analyzing the data quantitatively and qualitatively, including subjective analysis by observers, content analysis of videotape, screen analysis by evaluator, face impression by evaluator, results of questionnaire, and completion time by evaluators.

8. Summarizing the data and make comments on the result. (Adopted form Zimmerman et al., 1995, Hallahan, 2001, p. 227, and Rinder, 2012)

In conclusion, in the laboratory or a controlled environment, usability testing provides more valuable and reliable feedback to researchers. However, the traditional version of usability testing needs more funding, longer time, and may not provide feedback right away to the researchers. Alternatively, methods and modification advised by Krug (2006) may provide useful results quickly. This could allow the organization/designer to conduct more frequent testing through design process.

2.4.1 Heuristic Usability Testing

Heuristic evaluation has been introduced by Nielsen and Molich(1990). He is one of the most common and widely used methods for finding interface usability problems.

An evaluator who analyzes the usability of interface conducts the technique. According to Nielsen, 3 and 5 evaluators can be enough to find out majority of the problems with web site interfaces. The main idea behind HE is to have a small set of evaluators examining the interface and evaluate the interface with the usability principles.

"Nielsen stresses the importance of focusing heuristic evaluation on key criteria, rather than a litany of every possible problem" (Hallahan, 2001, p. 228).

According to Nielsen and Molich (1990) He is a valid approach however it is not easy to conduct it. Although they claim that five evaluators are enough and more evaluator will not discover significantly better results, Redish et al. (2002) claim that twelve evaluators will give better results.

Even though, He is inexpensive and can be performed fast, the results of HE could lead wrong analysis due to biases of the evaluator. Therefore, in order to get better result, five to twelve people should perform the evaluation (Hallahan, 2001, p. 228).

2.4.2 Other Frequently Used Usability Testing Methods

Card Sorting: A card sorting is one of the most popular ways in order to find out the structure of a web site. It gives an opportunity for users to decide which links or labels should be together in the web site, under what name these should be grouped. They organize set of items into groups, and then label the name of groups.

Persona Test: The purpose of personas is to write a real character with backgrounds, goals, and values; so that evaluators have a reliable and realistic representation of their key audience.

A/B Test: Putting 2 different design options for visitors who select which web site design option is best for them. Companies such as Google, Amazon, and EBay used this method many times for their web sites.

Surveys: They are an efficient and cost effective way in order to understand what users think about the web sites. A survey method is one the most used methods used for usability evaluation. Details regarding surveys will be explained in below.

Remote Evaluation: Popularity of remote evaluation always increases. Evaluators and participants can be in different places. Participants can choose the time and place

for their convenience. Participants are asked to complete some tasks. Provided systems can measure participants' task time completion, number of clicks, and other details.

User Testing: It is one of the most used usability measurement method. Participants attend a room which is designed with special equipments such as cameras, recording devices, special eye-glasses. Participants are asked to complete some tasks. Evaluators observed the whole process and take notes for measurements.

2.4.3 Combination of Usability Techniques

For all techniques, only one single evaluator is not satisfactorily enough to detect the majority of the usability. Therefore, some researchers combined several methods and proposed new one in order to obtain more accurate results such as combination of questionnaires, interviews, think aloud method, users testing, heuristics evaluation, and so on (Fernandez et al., 2011, Yussof et al., 2013, Waljia et al., 2014).

A study (Virzi et al., 1993) compared three methods and found that each method found almost equal number of problems and suggested heuristics and think aloud methods are more sensitive and uncovering more problems in user interface.

In Another study (Fu et al., 2002) user testing and heuristic evaluation, method has been compared. Weaknesses and strengths of both methods have been discussed.

2.5 RESEARCH TECHNIQUES REGARDING WEB SITE EVALUATION

There The necessity for evaluation of quality and usability of the web sites has been discussed on many articles (Oztekin et al., 2009, Elling et al., 2012). There are many techniques used to improve the usability of interactive systems. Some of these methods measured the usability of web sites using guidelines and heuristics while others used different techniques such as questionnaires and user testing. These techniques can be classified as follows: guidelines for improving system usability (Smith and Mosier, 1986), heuristics to evaluate user interfaces (Molich and Nielsen, 1990), methods to detect usability problems (Walji et al., 2014), discussions on how to measure usability (Nielsen and Levy, 1994, ISO, 1998, Hornbaek, 2006, Walji et al., 2014, Fogli and Guida, 2014) and checklists to evaluate usability (Oztekin, 2009, Lee and Kozar 2011, Elling et al., 2012). In addition, some researchers proposed hybrid

techniques combining several methods such as questionnaires, interviews, think aloud method, users testing, heuristics evaluation, and so on (Fernandez et al., 2011, Yussof et al., 2013, Walji et al., 2014).

2.6 COMPARISON OF WEB SITE METHODOLOGIES

The Most of the studies claimed that they developed some usability techniques but they have some validity and/or reliability issues (Muyllea et al, 2003, Oztekin et al, 2009, Elling et al., 2012). The majority of studies only collected data from using surveys without asking users' opinions for the websites. The feedbacks as qualitative data could uncover the majority of the usability problems where surveys only detect some of them (Walji, et al., 2014). Walji et al. (2014) claimed that survey only found 18% of the usability problems where user testing and interview discovered 54% and 24% of the problems respectively. Combining methods will increase the chance to uncover the usability problems (Marsico and Levialdi 2004, Walji et al., 2014).

In addition, there is confusion about the dimensions of usability and its definition. factors While satisfaction, learnability, and memorability has been defined under usability (i.e. ISO-9241, Nielsen, 1998), some researchers claim that satisfaction belongs to cognitive studies because it belongs to the emotional sphere, learnability and memorability are related with effectiveness because they enable users to complete their purposes (Fogli and Guida, 2014).

In this context, there is an emerging need for definition of usability and its factors. Additionally, a valid and empirically reliable method for web site usability evaluation is still needed.

Furthermore, most web site usability evaluation methods have some problems such as number of tasks which users were asked to complete while navigating on web sites before evaluating them. In order to uncover usability problems of a system, users have to spend more time on web sites to become familiar with them. In this way, the interaction between participants and system, as well as quality of response could be increased.

Since user testing uncovers more problems (Walji et al, 2014) compared to questionnaires, in order to increase the quality of questionnaire, participants' time spent on the web site should be increased. For instance, some authors only used a survey with

a few tasks (e.g., see Oztekin et al., 2009) and some other did not even clearly explain the number of the tasks (Delice and Gungor 2009, Belanche et al., 2012, Wu et al, 2014).

Taking these into account, in order to construct a new instrument, we have benefited from the System Usability Scale (SUS), (Brooke, 1996), the Website Analysis Measurement Inventory (WAMMI), (Kirakowski et al., 1998), the Website User Satisfaction Questionnaire (Muylle et al., 2004), Nielsen Heuristics (HE), (1994), ServQual (Parasuraman etl al., 1988), Web Service Quality (Web-ServQual), (Li, 2002), Website Evaluation Questionnaire (WEQ), (Elling et al., 2012) and Usability of Web-based Information Systems (UWIS), (Oztekin et al., 2009). Additionally, we considered some of quality and usability dimensions together because there are some studies claiming that some of usability and quality dimensions are the same (Oztekin et al., 2009).

The reason why we choose the abovementioned checklists and Nielsen Heuristics is because the System Usability Scale, WAMMI, the Website User Satisfaction Questionnaire, Nielsen Heuristics, ServQual, WebServQual are one of the most used, discussed or validated approaches (Elling et al., 2012, Oztekin et al., 2009). Additionally, UWIS and WEQ have been selected because these approaches are used for informational web sites and they are relatively new studies about usability assessment. Moreover, while WEQ focuses on the System Usability Scale, WAMMI, the Website User Satisfaction Questionnaire and some other checklists, UWIS benefits from the Nielsen Heuristics, ServQual, WebServQual and so on. We selected these two methods because of their comprehensive comparisons.

A questionnaire for web site evaluation should be available for general purpose and open for analysis to assess the quality of the web sites. Second, the purpose of the questionnaire should be clear. Third, a questionnaire should have clearly defined factors, which evaluate the usability of the web sites. Fourth, it is important that the scale constructed should be reliable and valid (Elling et al., 2012).

Many questionnaires are constructed according to these principles in order to be simple and cost-effective. In addition, these questionnaires are designed to be applied to various web sites and provide different results for usability evaluation.

The System Usability Scale (SUS) (Brooke, 1996) consists of ten items. Participants can express their opinions with five-point Likert scales. In SUS, there are two dimensions: usability with eight questions and learnability with two questions (Lewis and Sauro, 2009). The result of the evaluation is determined with a SUS score between 0 and 100. SUS score can be used to compare the web systems. SUS is a simple and short questionnaire and can be used even with small sample sizes (Tullis and Stetson 2004). However, SUS does not provide a thorough impression for the usability of a web site since it has limited number of questions. In addition, it is debatable that SUS can be applied to an informational website (Elling et al., 2012).

The second questionnaire, which is mentioned frequently in usability literature, is the Website Analysis Measurement Inventory (WAMMI). WAMMI has 20 questions with five-point Likert scales. There are five dimensions in WAMMI: attractiveness, controllability, efficiency, helpfulness, and learnability. WAMMI can be used for most of the web sites and it is suitable for benchmarking. However, there are some concerns about reliability and validity issues of WAMMI (Elling et al., 2012).

The third questionnaire is the Website User Satisfaction Questionnaire by Muylle et al. (2004). This questionnaire was developed for the evaluation of commercial websites. 60 questions have been tested with 837 website users. 11 sub-dimensions have been defined under 4 main dimensions. Reliability of the dimensions is between 0.74 and 0.89. This questionnaire is not designed for informational websites (Elling et al., 2009).

Fourth method for measuring service quality is ServQual approach. ServQual has been developed by Parasuraman et al., (1988) with 7 dimensions and used widely by many researchers. However, it has been criticized by many authors in terms of reliability and validity of the scale (Cronin and Taylor 1992, Teas 1993).

The fifth method is web version of ServQual, the web-based ServQual which has been developed with six dimensions and 28 checklist questions by Li et al. (2002). This approach does not evaluate web site quality with a quantitative model (Oztekin et al, 2009).

The sixth method is UWIS which has 7 dimensions with 24 questions. This method is design for informational web sites. However, this method does not evaluate the web sites with qualitative model. In addition, the number of tasks is insufficient.

Moreover, this model does not have some important dimensions of the usability. For example, UWIS does not have Aesthetics and Clarity and Visibility dimensions.

The seventh method is WEQ which has 7 dimensions with 24 questions as well. Even though this method was designed for informational web sites, it only focuses on quantitative results and does not have some important dimensions of the usability as well. For instance, WEQ does not have Responsiveness and Controllability dimensions. Moreover, this method does not clearly explain the number of tasks.

The eighth technique selected for usability evaluation is Nielsen's Heuristics (1994). It has 10 different suggestions about web site usability. Even though Nielsen's Heuristics is inexpensive, and can be performed fast, it has some drawbacks as well. First, it is not a questionnaire or a test and it does not provide a quantitative approach. In addition, the results of Nielsen's Heuristics could be misleading due to biases of the evaluators.

Among these abovementioned methods, none of them collects participants' opinions after questionnaires in order to get qualitative data from users. In addition, number of tasks is either not explained clearly or less than necessary because users do not spend time on the web site sufficiently in order to evaluate it. Moreover, some checklist items of these methods need to be improved because usability of web site is changing continuously. Therefore, some features of web sites are becoming more crucial for users such as Controllability, Clarity and Visibility, and Technology (as Efficiency). There are insufficient details in previous checklists regarding those dimensions. Furthermore, some important aspects of questionnaires have not been stated clearly; such as details of content validity, pilot study for reliability, and sample size calculation. Lastly, those questionnaires have not explained a guideline regarding how the instruments have been constructed systematically for web site usability evaluation.

Therefore, there is a need to construct a well-founded, comprehensive web site evaluation methodology which considers these abovementioned issues. In this study, we have developed a new questionnaire for user interface (QUIN) which will be described in detail below.

Table 2.4 Comparison of usability evaluation methods

the Website User Satisfaction Questionnaire	**WAMMI**	**the System Usability Scale (SUS)**	**ServQual (Parasuraman,1988) and WebSerQual (Li, 2002)**	**10 Usability Heuristics(Nielsen, 1994)**	**WEQ**	**UWIS**
Ease of use	Attractiveness	Usability	Tangibles	Match between system and the real world	Ease of use	Navigation
Entry guidance	Helpfulness	Learnability	Reliability	Visibility of system status	Hyperlinks	Integration of communication
Structure	Efficiency		Responsiveness	User control and freedom	Structure	Controllability
Hyperlink connotation	Learnability		Assurance	Help and documentation	Comprehension	Responsiveness
Speed	Controllability		Empathy	Consistency and standards	Completeness	Assurance
Relevance			Quality of information	Aesthetic and minimalist design	Lay out	Reliability
Accuracy			Integration of communication	Flexibility and efficiency of use	Relevance	Quality of information
Comprehensibility			Call-back systems	Error prevention	Search option	
Completeness			Web	Help users recover from errors		
Layout			Competence	Recognition rather than recall		
Language						

2.7 QUESTIONNAIRES

Questionnaire is one of the most used, traditional, cost-effective, and efficient way to collect data for usability evaluation (Ozok, 2008, Elling et al., 2012). Questionnaires can be used with other evaluation methods as well. When participants are being observed during a user test, they could complete some tasks and answer survey questions. They may provide great results if they are applied correctly.

However, surveys have some potential problems while collecting data that could affect their quality. These problems can be classified as less control over the survey and participants due to high number of participation and undesired participants, the non-response error (unwillingness to participate in the survey), sampling error (not all users have the same chance when starting the survey such as using different machines, smart phones or laptops) (Couper,2000). In addition, lacks of user attention, selection of sample, concern on privacy are other problems regarding surveys. Moreover, the accuracy of the answers, scale to decide the level of opinions (could be perceived differently by each participant), expected benefit and perception on the importance of survey can be mentioned as other concerns on quality of surveys (Heerwegh and Loosveldt, 2008; Galesic and Bosjnak 2009; Couper, 2000, Tourangeau et al., 2000; Tourangeau, 2003, Elling et al., 2012). Furthermore, surveys may not find all usability problems (Walji et al., 2014). Therefore, there is a need to improve the quality of surveys for usability evaluations.

Some steps are needed to be taken to avoid the abovementioned concerns. In this study, the participants completed a group of tasks regarding the web site, then they responded to a set of questions. Navigating across the web site, spending time to complete the tasks result in more reliable responses. Secondly, participants were invited by email for heterogeneity and to avoid undesired participation. Participants completing the survey were from a variety of disciplines such as engineering, management, theology, art and science, prep school and so on. Participants are asked to take survey through computer in order to prevent the sampling errors. Thirdly, teachers have explained the importance and benefits of the survey in class. Fourthly, participants have been rewarded with grades to increase the participation and in order to avoid the non-response rate. Additionally, it has been

mentioned at the top of the survey that participants privacy will be protected; their name and response will not be shared with anyone. Moreover, a brief guideline before survey has been stated to clarify how to put their valuable opinions into the survey; the scale for determining level of opinion. Finally, time flexibility for completing survey has been offered; they have been told to complete the questionnaire whenever/wherever they want to so participants would spend their time and express their opinions in natural environment without any pressure (Spyridakis et al., 2005).

CHAPTER 3

INSTRUMENT CONSTRUCTION

In this chapter, a methodology will be described regarding how a new instrument can be constructed.

3.1 CONSTRUCTION OF INSTRUMENT FOR USER INTERFACE

Before constructing an instrument, there is a need to define a guideline for checklist construction process. Figure 3.1 shows how the instrument has been constructed systematically.

The process for the instrument construction has been started with defining factors for usability and their corresponding questions by reviewing literature. After detailed literature review for usability as seen in Table 3.1, dimensions of the QUIN instrument, which have been mentioned from different checklists or questionnaires, have been selected. Afterwards, a case study has been applied in order to collect data. Finally, reliability and validity of the instrument has been checked as seen in Figure 3.1.

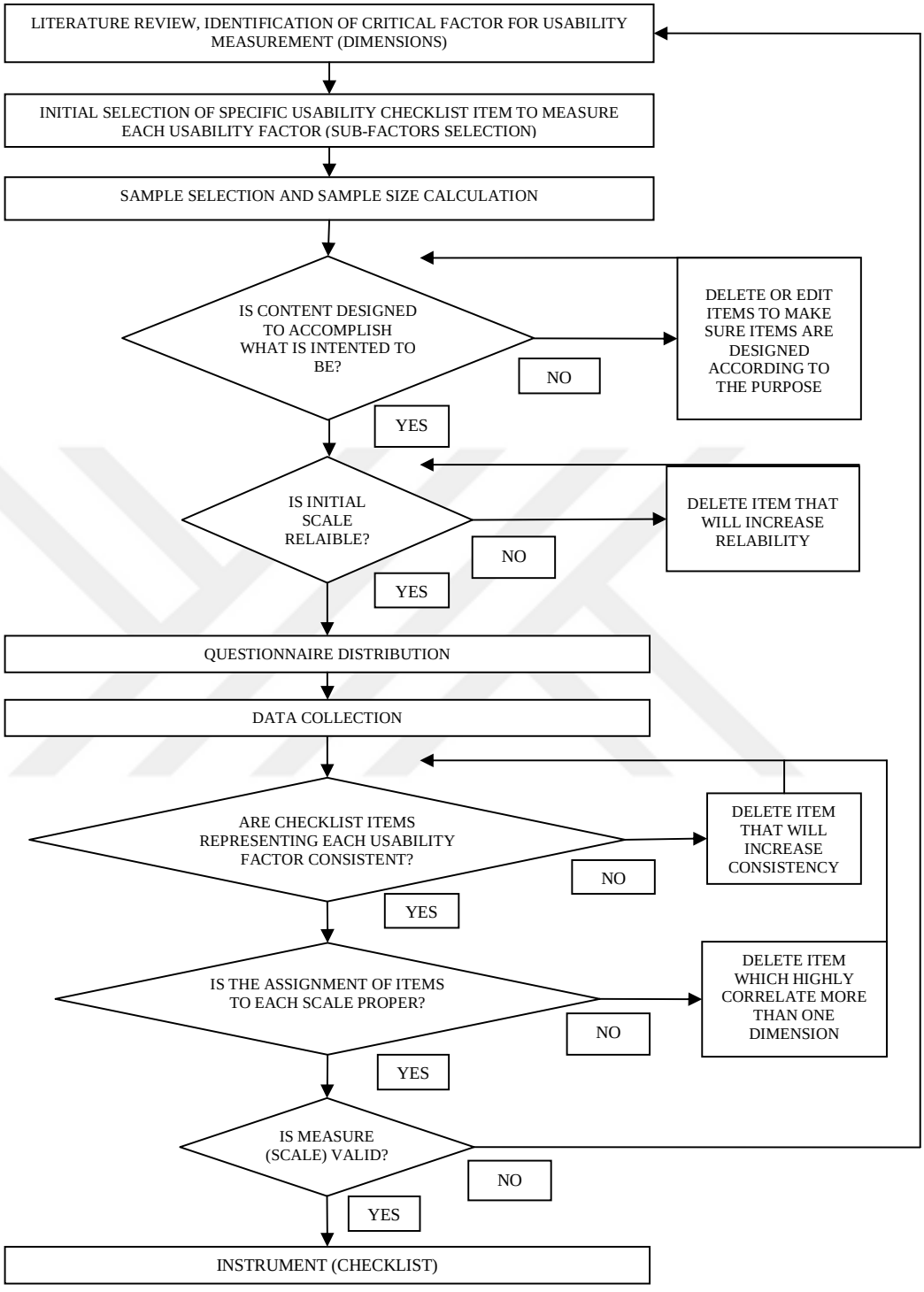


Figure 3.1 The Process of QUIN Construction (adopted form Saraph et al. 1989).

Table 3.1 Comparison of QUIN method with other checklists.

the Website User Satisfaction Questionnaire	**WAMMI**	**the System Usability Scale (SUS)**	**ServQual (Parasuraman,1988) and WebSerQual (Li, 2002)**	**10 Usability Heuristics(Nielsen, 1994)**	**WEQ**	**UWIS**	**QUIN**
Ease of use	Attractiveness	Usability	Tangibles	Match between system and the real world	Ease of use	Navigation	Easiness
Entry guidance	Helpfulness	Learnability	Reliability	Visibility of system status	Hyperlinks	Integration of communication	Clarity and Visibility
Structure	Efficiency		Responsiveness	User control and freedom	Structure	Controllability	Controllability
Hyperlink connotation	Learnability		Assurance	Help and documentation	Comprehension	Responsiveness	Responsiveness and Help
Speed	Controllability		Empathy	Consistency and standards	Completeness	Assurance	Completeness(System Assurance and Reliability)
Relevance			Quality of information	Aesthetic and minimalist design	Lay out	Reliability	Aesthetic (layout)
Accuracy			Integration of communication	Flexibility and efficiency of use	Relevance	Quality of information	Efficiency (Speed)
Comprehensibility			Call-back systems	Error prevention	Search option		Information Quality
Completeness			Web	Help users recover from errors			Search
Layout			Competence	Recognition rather than recall			
Language							

3.2 UNIQUE FEATURES OF QUIN METHODOLOGY

The goal of this study was to present a novel checklist. Therefore, QUIN methodology was aimed to fulfill the gap regarding definition of the usability factors for the web sites. Since, still there is no consensus on the factors and evaluation methods of usability (Oztekin et al, 2009, Fogli and Guida, 2014), a unique checklist was needed to be constructed. There are six remarkable features of the QUIN methodology as following:

Firstly, QUIN has two primary data collection sections: quantitative data as survey questions and qualitative data as user feedbacks. The methodology is designed with respect to most valid and popular usability evaluation methods, checklist and guidelines. Supporting the quantitative data with qualitative findings has not been studied widely in literature because using only a survey without user feedbacks cannot find all usability problems.

Second remarkable aspect of the QUIN is the extra dimensions, which have not been well-mentioned among these checklists and guidelines such as efficiency, controllability, and clarity and visibility. Those dimensions have been added in QUIN instrument.

Third important feature of the methodology is the number of tasks and their details. QUIN has 6 tasks which have been explained clearly. Most of the previous studies do not explicitly state number of tasks and their details. Users need to spend time on web sites in order to experience usability of the systems (Elling et al, 2012). For instance, giving a few tasks to users and asking them to respond to many questions would not produce reliable results because users can achieve these tasks easily without examining all features of the web sites or complete the tasks by using only search option without navigating the web site. Increasing the number of tasks will enforce the participants to experience the web site.

The fourth remarkable feature of the methodology is the tasks which have been spread across the survey because in this way, users could experience the web site. For example, when users start the survey, they are asked to complete 2 tasks and answer 6 questions. Afterwards, participants are asked to complete 2 more tasks and respond to

10 more questions. Finally, they are asked to complete last 2 tasks and respond to last 13 questions.

The fifth remarkable feature of the methodology is the number of participants which has been decided according to Yamane's formula (1967). Required number of participants is calculated in order to increase the quality of sample. Insufficient participation may not reflect the population correctly; also may produce wrong results which could decrease the quality of the survey analysis.

Final important feature of the QUIN is the expert evaluation for content validity, pilot study for reliability, which has not been clearly explained in most of the previous studies. In other words, the number of experts for content validity and number of participant in pilot study for the reliability of instrument have been stated clearly in QUIN methodology.

For instance, QUIN includes all WAMMI factors, one of the most valid and widely used questionnaires. As seen in Figure 3.3, QUIN includes all dimensions of the WAMMI, also QUIN has more dimensions, which are important for users and cannot be ignored. In addition, QUIN covers most of Nielsen Heuristics and ServQual method. Moreover, looking at the other methods at Table 3.1, it can be seen how QUIN covers almost all of the methods such as SUS and WEQ.

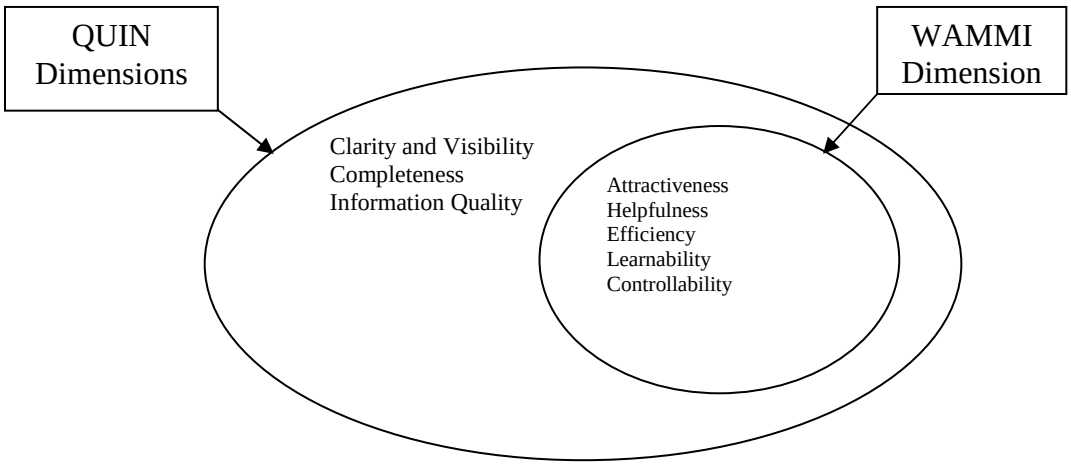


Figure 3.2 Dimension of QUIN compared to WAMMI

Additionally, QUIN is that it includes the perspective that is more comprehensive. For example, QUIN includes Efficiency (as speed), Easiness (includes learnability and memorability), and Aesthetics (as layout), Controllability (includes personalization, customization) even though some of them are mentioned partly/separately. Almost none of the abovementioned methods include all dimension/factors of QUIN, which uncover the real usability problems. Furthermore, QUIN is the feedbacks collected from users, which uncovers the usability problems from different perspective. Therefore, it can be stated that QUIN with extra factors and feedback support has a potential to bring a new perspective for evaluating usability of the informational web sites.

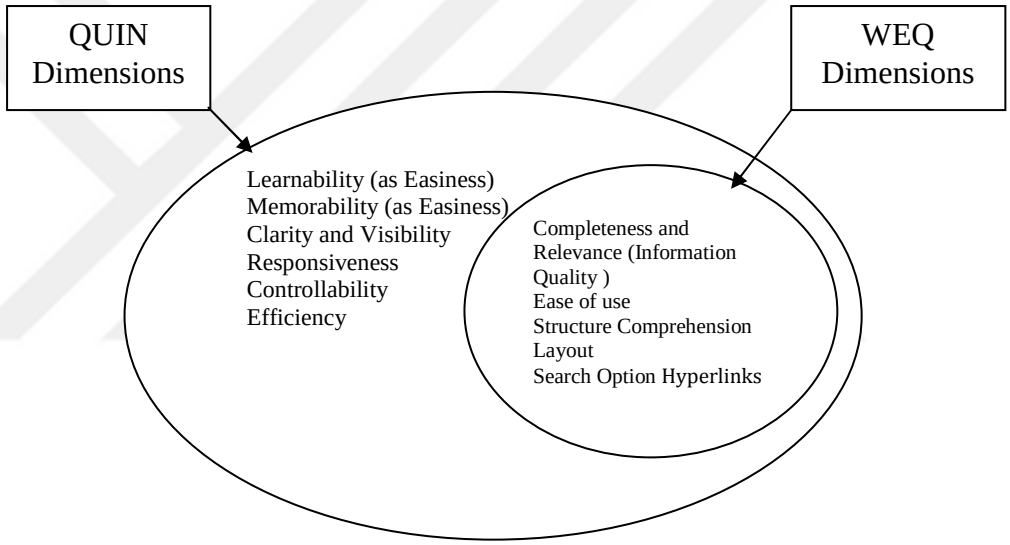


Figure 3.3 Dimension of QUIN compared to WEQ

Completeness and Relevance in WEQ is almost the same with Information Quality Dimension in QUIN. In addition, Ease of use and Structure and Comprehension dimensions in WEQ is almost the same with Easiness dimension in QUIN. However, Easiness dimension in QUIN has extra dimensions as learnability and memorabality. QUIN had Search Option however; it has been deleted after factor reduction because Search Option in the student portal of the university was not working properly which caused inconsistent data for SPSS. Hyperlinks in WEQ are almost the same as Completeness dimension that includes system assurance, reliability in QUIN.

QUIN dimensions have been defined based on 29 equally distributed questions from the quality, and usability assessment approaches. In addition, at the end of QUIN questionnaire, participants' opinions have been collected. Qualitative data from the participants such as what they liked or disliked about the web site when completing the tasks and responding to the related question have been gathered. In addition, as explained before, QUIN questionnaire includes six tasks, which have been divided into three parts. In this way, participants have really experienced the web site and answered questions. The number of tasks was designed to make sure that users really experience the web site, which is a plus for QUIN compared to other questionnaires.

Moreover, QUIN has been developed as an online survey, which is more usable, and cost-effective option for data collecting compared to paper based questionnaire. In addition, QUIN has been distributed to actual users because only participants who have student id could enter the student portal to complete tasks. This consideration is one of the most important issues in online surveys (Elling et al, 2012).

3.3 QUIN MEASUREMENT MODEL

In QUIN evaluation method, the questions (observed variables) load on dimensions (latent variables). First order latent factors compose the second order factor namely usability. The purpose of the structure is to see how factors explain usability. In other words, the question "how many factors does usability has?" is tried to be answered. In addition, each factor weight has been investigated in this method. Each first latent factor would be measured with an observed variable and the second order latent variable would be measured with first latent factors. We used AMOS 22 in order to measure the weights between second order latent factor, first order latent factors and observed variables.

Furthermore, user feedbacks will be collected at the end of each questionnaire. User feedbacks will be compared with questionnaire responds. This enriched comparison will help to define real usability problems and factors.

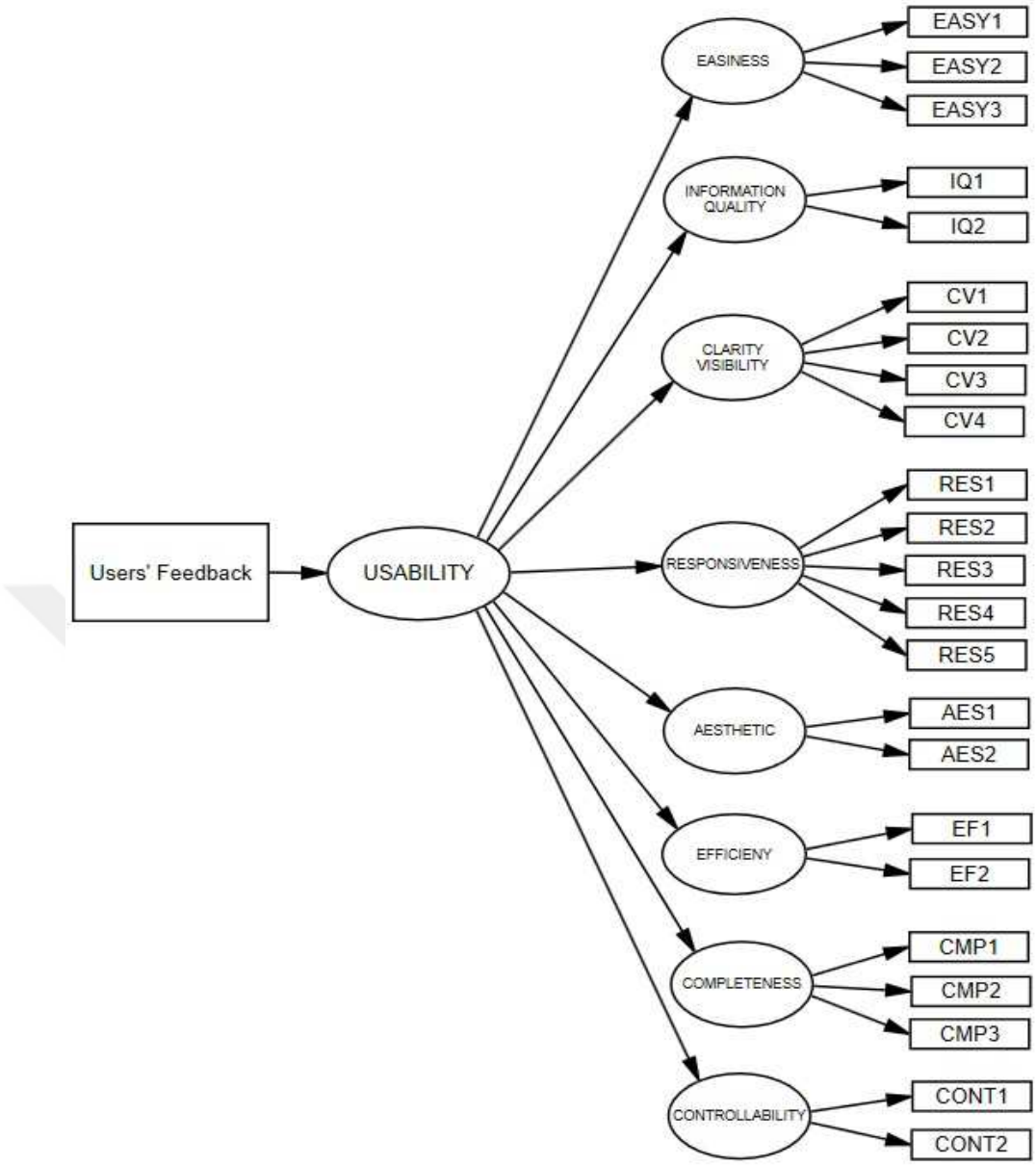


Figure 3.4 The Structure of QUIN model.

The questionnaire is designed to collect two types of data: a quantitative evaluation with 29 questions with 5 point Likert scale (23 final version of question), and open-ended field where participants express their opinions freely.

The goal of the study is to present a new web site usability evaluation approach to literature involving users' perspectives which practitioners and scholars can utilize while designing web sites.

Steps in the research were:

1. Introducing and defining a web site usability evaluation instrument;

2. Identifying the underlying dimensions of the usability checklist and sub factors;

3. Empirically validating the instrument;

4. Supporting the instrument with user feedbacks;

5. Determining critical usability problems considering quantitative and qualitative results;

6. Offering solutions to the problems;

7. Giving recommendations for future studies.

CHAPTER 4

IMPLEMENTATION

4.1 DISTRIBUTION PROCESS OF THE QUESTIONNAIRE

In this chapter, the reliability and validity of methodology will be reviewed. Since QUIN methodology can be applied to informational web sites, a university web portal was selected as case study. Some important steps will be explained before distributing the survey to the participants.

4.1.1 Sample Size Calculation

In order to determine the sample size required, Yamane's formula (1967) has been used. For 95% confidence level, the required sample size is calculated as follows:

$$n = \frac{N}{(1 + N\,e^2)}$$

n = sample size required

N = total population (13929)

e = error tolerance (5%)

= 387.75

= 388 required participants.

Even though 415 responses for the survey invitation have been obtained, 26 participants dropped out somewhere in the survey. In addition, not all participants responded to the survey seriously. Thus, 40 participants (9.7%) were deleted from dataset. Moreover, whoever responded to 25 out of 29 questions using the same scale

was excluded from the survey (Elling, et al., 2012). In that way, we eliminated 43 more participants and finally had 346 respondents which were slightly lower than the desired number. However, according to the formula, required sample sizes with corresponding error tolerances can be listed as follows:

Table 4.1 Required sample sizes with their corresponding error tolerances.

E	n
0.01	5821
0.02	2120
0.03	1029
0.04	598
0.05	389
0.0531	**346**
0.06	272
0.07	201
0.08	155
0.09	122
0.1	99

As seen in Table 4.1, even though required sample size is 388 for 95% confidence level, the number reached (i.e. 346) has 5.31 % error tolerance corresponding to 94.69% confidence level, which is highly acceptable. In addition, according to Rigdon(1998), 5 or 10 observations per survey question would be satisfying enough while we had 346 people for 29 questions.

4.1.2 Content Validity Before Distribution

Before distributing the questionnaire, content validity has been checked in order to make sure whether the questionnaire is appropriately designed for what it is intended to be with the help of two researchers who have publications on usability. In addition, four faculty members reviewed the questions in QUIN and suggested some corrections.

4.1.3 Reliability Test

After determining the sample size, a pilot study has been conducted to check the reliability of the instrument. While Isaac and Michael (1995) and Hill (1998) suggested 10–30 participants would be enough, Julious (2005) claimed that 12 participants would be ideal for the reliability of pilot study. In this aper, the pilot study with 20 students has been performed before distributing the survey.

For reliability of pilot study, IBM SPSS 20 statistical software is used. Cronbach's alpha value has been obtained as 0.914 for the reliability of the scale, which confirms internal consistency.

4.1.4 Construct Validity of the Questionnaire

According to Cao and Dowlasthahi (2005), there are three ways to assess the construct validity: unidimensionality, discriminant, and convergent validity.

Unidimensionality gives a proof that a dimension is a single latent construct. There are two ways to find out unidimensionality of a measure: Exploratory factor analysis (EFA) and Confirmatory Factor analysis (CFA). In this study, EFA was used to explore the dimensions of the usability.

Discriminant validity refers that different constructs should be dissimilar (Burns and Bush, 1995). In order to check discriminant validity between constructs, we used the Average Variance Extracted (AVE) and the squared-correlation between two constructs. According to Fornell and Larcker (1981), in order to assess the discriminant validity, square-root of AVE for a construct should be greater than the correlations involving the construct. Moreover, AVE for each construct should be greater than 0.50 (Fornell and Larcker, 1981).

Convergent validity can be tested by factor loading, Composite Reliability (CR) and AVE (Fornell and Larcker, 1981). An acceptable factor loading value should be greater than 0.5. However, if it is equal to 0.7 or above, it is considered good.

Secondly, CR is another way to check convergent validity. CR measures the level of variance captured by a construct versus the level due to measurement error. The acceptable value of CR is 0.7 and above (Hair et al., 2010).

CR can be calculated as follows:

$$CR = \frac{(\sum_{i=1}^{n} \lambda_{yi})^2}{(\sum_{i=1}^{n} \lambda_{yi})^2 + (\sum_{i=1}^{p} Var(\varepsilon_i))}$$

CR : Indicates composite reliability

λ_y : The standardized factor loading

$Var(\varepsilon_i)$ = The variance due to the measurement error

The third method to check convergent validity is Average Variance Extracted (AVE). If AVE is more than 0.5, it is considered acceptable. (Hair et al., 2010). It is calculated as:

$$AVE = \frac{\sum_{i=1}^{n} \lambda_i^2}{n}$$

AVE: Average variance extract

λ_i : The standardized factor loading

n : The number of items on a dimension

4.3 DATA COLLECTION

In this study, we utilized the QUIN instrument. We aimed to collect both quantitative and qualitative data as seen in Figure 5. The process for quantitative data collecting was survey questions. In addition, we collected qualitative data asking the opinions of the participants at the end of each questionnaire. QUIN is composed of five sections; first section is background information, successive three sections include the questionnaire and six different tasks, using 5 point Likert scales (i.e. 1.Strongly Disagree; 2. Disagree; 3.Neutral; 4. Agree; 5. Strongly Agree), and last part is a field where users freely comment on the problematic issues of the web portal.

The study is applied to 2015-2016 academic year. In order to define the demographic factors, some question has been asked as seen in Table 4.2.

Table 4.2 Demographics of participants.

Factors	Category	%
Gender	Male	51
	Female	49
Grade	First	16
	Second	30
	Third	24
	Fourth	26
	Graduate	3
Weekly computer usage online	Never	0
	1-6 hours	52
	>=7 hours	48
Myfatih Visits Weekly	Never	5
	1- 6 times	67
	>=7 times	28

Investigating the demographics factors, we can see that the percentage of male and female is almost equal. Additionally, the percentages of respondents' classes are close to each other except the percentage of graduate students. A remarkable outcome is high proportion of the participants, 28%, are visiting university student portal more than 7 times a week and 52% of users using the portal 1 to 6 times a week. Thus, most of the users are very familiar to system. Throughout the questionnaire, participants are asked to complete some tasks. The reason behind this idea is to make sure participants are responding to the questions with the experience they just had while accomplishing the tasks. Otherwise, participants would respond to the questions with their past-experience, which could generate inadequate results. Tasks were as follows:

Task 1: Please find out how many hours you are available on Thursday on your school schedule.

Task 2: Please read the last message coming from the university and mark it as "read".

Task 3: Please help one of your foreign friends who asked you to help him/her to change the language of the system to English.

Task 4: Please find the details of a course you take this year.

Task 5: Please find the place in the system where your GPA is recorded.

Task 6: Please make a reservation for Thursday at the fitness room.

4.4 RESULT AND DISCUSSIONS

The descriptive statistics in Table 4.2 show mean and standard deviation for each checklist item in the QUIN. From Table 4.2, the question *"The web site can be personalized for each visitor"* has the lowest mean explaining that participants disagree with content of the question. In addition, the question *"The wording of the website is clear and visible"* has the highest mean explaining that participants agree with the content with the question.

Table 4.3 The result of descriptive statistics.

Checklist Items in QUIN	Mean	SD
I can reach the help documents in case of necessity.	3.26	.09
This website provides clear and useful messages if I don't know how to proceed.	3.22	.05
Web page always provides me to return to first or previous page at any time (e.g. links to homepage) in case of necessity.	3.56	.03
Messages of error or warning are understandable and prevent possible errors from reoccurring.	3.41	.11
There are security message and e-mail notification on the system.	3.35	.11
It does not take too much time to find information needed when clicking through on the website.	3.79	.80
I don't see any difficulty to understand in this website.	3.95	.85
I remember easily everything on the website and I know where to find them	3.93	.85
The information on the web site enough and helpful.	3.86	.84
I can rely on the information on the web site and it is up-to-date.	4.07	.71
It is not difficult to read the number on the web site.	4.02	.97
The wording of the website is clear and visible.	4.14	.93
The harmony of colors and structures on the website are great for the eyes.	4.01	.90
The interfaces include the title of the site and the sections in a visible way.	3.86	.90
The website contains similar name, the same structure across the pages.	3.82	.79
Pages, titles, and links have been gathered in order (Clustering, grouping and sub-links).	3.82	.90
The activity icons, buttons, labels, and links are designed with respect to the purpose intended.	3.82	.91
The web site is quick because it loads pages fast.	3.37	.27
The web site designed with respect to technology.	3.10	.14
The web site can be personalized for each visitor.	2.68	.29
It is possible to increase and to reduce the size of the font.	3.01	.25
The design of interface on the website is attractive and it is pleasing to look at screens.	3.56	.96
I feel comfortable with the colors and graphical illustrations used on the web site (e.g. colors, pictures).	3.51	.05

Even though we had 29 questions initially, we only mentioned 23 questions here, the final version of the questionnaire. We will explain the details in exploratory and confirmatory factor analysis.

4.4.1 Exploratory Factor Analysis (EFA)

EFA was conducted with 29 questions in order to find the usability dimensions. As explained before, these questions were selected carefully from the usability guidelines and checklists. Even though, questions have been selected from literature, it has been explored to see how many factors belong to usability for informational web sites. After deciding the questions for dimensions, EFA was conducted. IBM SPSS 20 was used in order to analyze the statistical calculations. 29 questions have been added into the SPSS for data reduction. When QUIN instrument was designed, a systematic way has been followed to eliminate the items from the instrument. For instance, items loading has been checked; if an item is loading less than 0.30, it has been eliminated (Hair et al., 1995, De Vaus, 2001).

After eliminating the items from the checklist carefully, 6 factors and 23 checklist items has been obtained. Search dimension was deleted completely and some other questions in the questionnaire. Additionally, the reliability of the data was checked. Reliability test is used to see whether the result from the data analysis is replicable. For all items, Cronbach's Alpha value which is used for the initial reliability of a test or scale was checked (Cronbach, 1970). It is expressed as a number between 0 and 1. The rule is expressed as follows: if the value "> .9 – Excellent, > .8 – Good, > .7 – Acceptable, > .6 – Questionable, > .5 – Poor and < .5 – Unacceptable" (George and Mallery, 2003). Therefore, reliability of the data set is 0.898, which is excellent.

For factor analysis and data reduction, Varimax orthogonal rotation is used. In addition, Kaiser normalization is used for extraction, and its default Eigenvalue was set to 1. In spite of the fact that our data has excellent reliability, Efficiency and Information Quality dimensions have been deleted based on Eigenvalues. According to Eigenvalue analysis, in order to have a factor, its Eigenvalue must be greater than 1 (Kaiser, 1960). As seen in Table 4.4, Eigenvalue of the 6 factors are greater than 1.

Table 4.4 Exploratory factor analysis of QUIN with 6 factors.

Symbol	Variables	Factors					
		1	2	3	4	5	6
RES1	I can reach the help documents in case of necessity	.776					
RES2	This website provides clear and useful messages if I don't know how to proceed.	.765					
RES3	Web page always provides me to return to first or previous page at any time (e.g. links to homepage) in case of necessity.	.789					
RES4	Messages of error or warning are understandable and prevent possible errors from reoccurring.	.720					
RES5	There are security message and e-mail notification on the system.	0.8					
EASY1	It does not take too much time to find information needed when clicking through on the website.		.657				
EASY2	I don't see any difficulty to understand in this website.		.687				
EASY3	I remember easily everything on the website and I know where to find them		.673				
IQ1	The information on the web site enough and helpful.		0.7				
IQ2	I can rely on the information on the web site and it is up-to-date.		0.65				
CV1	It is not difficult to read the number on the web site.			0.82			
CV2	The wording of the website is clear and visible.			0.83			
CV3	The harmony of colors and structures on the website are great for the eyes.			0.76			
CV4	The interfaces include the title of the site and the sections in a visible way			0.65			
CMP1	The website contains similar name, the same structure across the pages.				0.6		
CMP2	Pages, titles, and links have been gathered in order(Clustering, grouping and sub-links)				0.62		
CMP3	The activity icons, buttons, labels, and links are designed with respect to the purpose intended.				0.62		
EF1	The web site is quick because it loads pages fast.				0.71		
EF2	The web site designed with respect to technology.				0.67		
CONT1	The web site can be personalized for each visitor.					0.83	
CONT2	It is possible to increase and to reduce the size of the font.					0.82	
AES1	The design of interface on the website is attractive and it is pleasing to look at screens.						0.83
AES2	I feel comfortable with the colors and graphical illustrations used on the web site (e.g. colors, pictures).						0.83

The factors have been named as Content Easiness, Aesthetics, Clarity and Visibility, Responsiveness and Help, Completeness, and Controllability for 6 factors.

Table 4.5 Reliability statistics with 6 factors.

Dimensions	Cronbach's Alpha
Content Easiness	0.808
Aesthetics	0.68
Clarity&Visibility	0.855
Responsiveness	0.879
Completeness	0.812
Controllability	0.772

Even though Cronbach Alpha value for Aesthetics is 0.68, it is close to the margin for acceptance value (0.7). Other factors are in acceptable range. Therefore, all factors (except Aesthetics) support the construct reliability.

Even though six factors have been defined with respect to Eigenvalues according to Kaiser's suggestion, there are some critics regarding Eigenvalues (Bhattacharya, 2015, Jolliffe, 1972, .and Hair et al., 2010). For instance, Bhattacharya (2015) used 0.9 while Jolliffe(1972) suggested that even 0.7 could be enough for Eigenvalues. When we checked the Eigenvalues with default set up value of 1 compared to 0.7, we saw that Kaiser's criterion is too strict because it deletes some of the dimensions. Moreover, default Eigenvalue of 1, could produce wrong result by over-extracting factors because some important factors can be dropped from the model (Fabrigar et al., 1999, Gorsuch, 1983). Therefore, we also checked Eigenvalues with 0.7 in order to see the differences.

After setting the Eigenvalue to 0.7, we obtained 8 factors explaining the 74.7% of the observed variance which is 7% more compared to result of six factors' observed variance. In addition, 23 items loaded on eight factors are shown on Table 4.6. As can be seen, the new model incorporates the Information Quality and Efficiency dimensions.

In addition, the Kaiser-Meyer-Olkin measure of sample adequacy is found as 0.898 which is almost excellent (Field, 2009). We named the factors of the proposed model as Easiness, Information Quality, Aesthetics, Clarity and Visibility, Responsiveness and Help, Completeness, Efficiency, and Controllability. If we look at the Cronbach's Alphas for 8 factors, all factors except Aesthetics are in acceptance range and support the construct reliability; Aesthetics' value is 0.680 which is very close to the margin of 0.7. The values of Cronbach's Alpha for each factor are 0.780, 0.743, 0.680, 0.855, 0.879, 0.801, 0.815, and 0.772 respectively.

Table 4.6 Exploratory factor analysis of QUIN with 8 factors.

Symbol	Variables	Factors							
		1	2	3	4	5	6	7	8
RES1	I can reach the help documents in case of necessity	.773							
RES2	This website provides clear and useful messages if I don't know how to proceed.	.759							
RES3	Web page always provides me to return to first or previous page at any time (e.g. links to homepage) in case of necessity.	.795							
RES4	Messages of error or warning are understandable and prevent possible errors from reoccurring.	.709							
RES5	There are security message and e-mail notification on the system.	.792							
EASY1	It does not take too much time to find information needed when clicking through on the website.		.819						
EASY2	I don't see any difficulty to understand in this website.		.814						
EASY3	I remember easily everything on the website and I know where to find them		.777						
IQ1	The information on the web site enough and helpful.		.656						
IQ2	I can rely on the information on the web site and it is up-to-date.			.679					
CV1	It is not difficult to read the number on the web site.			.753					
CV2	The wording of the website is clear and visible.			.749					
CV3	The harmony of colors and structures on the website are great for the eyes.				.691				
CV4	The interfaces include the title of the site and the sections in a visible way				.753				
CMP1	The website contains similar name, the same structure across the pages.				.745				
CMP2	Pages, titles, and links have been gathered in order(Clustering, grouping and sub-links)					.850			
CMP3	The activity icons, buttons, labels, and links are designed with respect to the purpose intended.					.839			
EF1	The web site is quick because it loads pages fast.						0.8		
EF2	The web site designed with respect to technology.						0.8		
CONT1	The web site can be personalized for each visitor.							.786	
CONT2	It is possible to increase and to reduce the size of the font.							.797	
AES1	The design of interface on the website is attractive and it is pleasing to look at screens.								.834
AES2	I feel comfortable with the colors and graphical illustrations used on the web site (e.g. colors, pictures).								.827

4.4.2 Confirmatory Factor Analysis (CFA)

CFA was used to test EFA results in the model to see whether the model fits the data (Bandalos, 1996). CFA usually tests an existing model. The observed data in EFA and hypothesized CFA model are evaluated with some fit statistics.

For confirmatory factor Analysis, we used Amos 22. Using the results of CFA, we can analyze how model represents the data well enough. For model fit criteria, the researchers suggest that using multiple criteria to evaluate the model (Fan et al., 1999) such as chi-square (χ^2/df), goodness of- fit index (GFI), adjusted goodness-of-fit index (AGFI), and root mean square residual (RMS) are usually used (Schumacker and Lomax, 1996).

The model fit statistically can be assessed by Chi-square (Sharma, 1996). In large samples the index Normed by Chi-square is suggested which is χ^2/df. Values between 1 and 3 seem good fit (Hair et al., 1995). It is suggested that using χ^2/df, GFI, AGFI,

RMR, and CFI values together are important for judging the model fit (Satorra and Bentler, 1994).

The goodness-of index called GFI represents the overall degree of fit (Hair et al., 1995). GFI is suggested to be greater than 0.90 (Joreskog and Sorbom, 1993). Comparative fit index called CFI represents the comparison between the estimated model and null model and if CFI is greater than 0.90, it is considered acceptable (Joreskog and Sorbom, 1993). Adjusted goodness-of-fit (AGFI) is also suggested to be closer to 1 (Hooper et al., 2008).

The root mean square error of approximation (RMSEA) value from 0.05 to 0.08 is acceptable (Hair et al., 1995). RMR root mean square residual (RMR) value is suggested to be less than 0.05 (Hair et al., 2010).

Amos 22 was used to test the validity of the model. Selected data from survey has been added into the program. We have run the model for both 6 and 8 factors to compare the validity results.

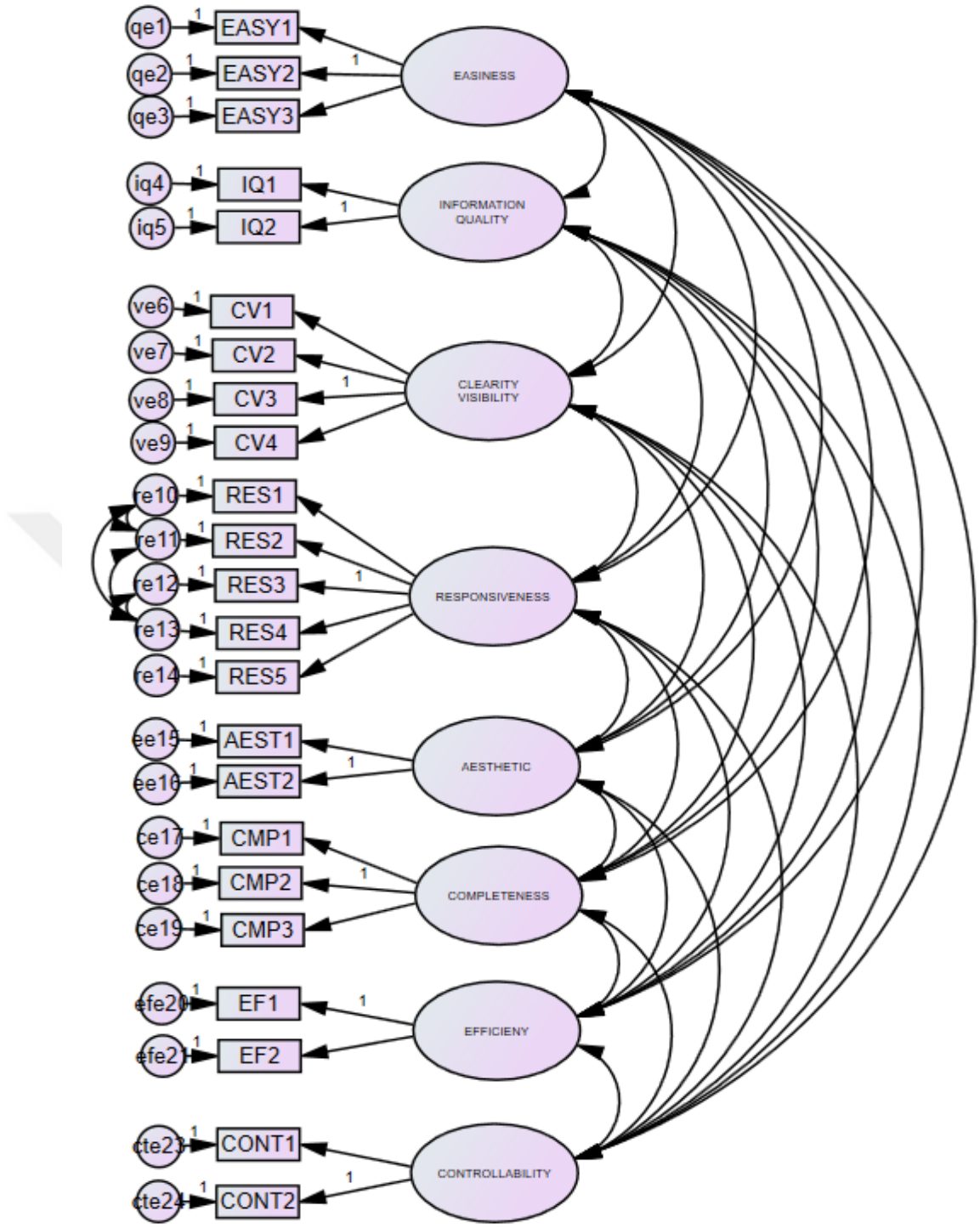


Figure 4.1 Model presentations in Amos.

If both models have been run in Amos, the results can be obtained as seen in the Table 4.7

. Table 4.7 Goodness-of-fit statistics (N = 346).

Number of Factors	χ^2	p-Value	χ^2/df	RRMR	GFI	AGFI	CFI	RMSEA
For 6 Factors	410.435	0	.032	0.067	.908	.874	.944	0.055
For 8 Factors	339.087	0	.713	0.05	.923	.892	.962	0.045

We can see that 6 factors model is almost fit except for RMR. However, model fit better with 8 factors including RMR. Also, all statistical values with 8 factors are better than 6 factor model. We can say that result for both model support the construct validity for the QUIN measurement model but 8 factors l is better. Namely, the model fit for both model is sufficient enough but 8 factors model has better values.

Considering Eigenvalue restriction which deletes important factors in 6 factor model (if those dimensions have not been deleted they caused validity issue in validation process) and CFA result for both models, 8 factor model has been chosen. Therefore, convergent and discriminant validity for 8 factor model will be demonstrated here.

Table 4.8 also presents all checklist items in QUIN with their corresponding loadings and their t Value. All loading values are statistically significant for the model because their t-values are not in the range (-1.96 +1.96).

Table 4.8 Confirmatory factor analysis.

Symbol	Description	Regression weight	t Value
COMPLETENESS			
CMP1	The website contains similar name, the same structure across the pages.	0.714	13.03
CMP2	Pages, titles, and links have been gathered in order(Clustering, grouping and sub-links)	0.795	-
CMP3	The activity icons, buttons, labels, and links are designed with respect to the purpose intended.	0.769	14.06
EFFICIENCY			
EF1	The web site is quick because it loads the pages fast.	0.837	-
EF2	The web site designed with respect to technology.	0.826	14.18
CONTROLLABILITY			
CONT1	The web site can be personalized for each visitor.	0.797	10.02
CONT2	It is possible to increase and to reduce the size of the font.	0.79	-
CLARITY& VISIBILITY			
CV1	It is not difficult to read the number on the web site.	0.817	14.43
CV2	The wording of the website is clear and visible.	0.829	14.61
CV3	The harmony of colors and structures on the website are great for the eyes.	0.737	-
CV4	The interfaces include the title of the site and the sections in a visible way	0.714	12.66
EASINESS			
EASY1	It does not take too much time to find information needed when clicking through on the website.	0.719	12.47
EASY2	I don't see any difficulty to understand in this website.	0.778	-
EASY3	I remember easily everything on the website and I know where to find them	0.714	12.38
INFORMATION QUALITY			
IQ1	The information on the web site enough and helpful.	0.781	11.7
IQ2	I can rely on the information on the web site and it is up-to-date.	0.768	-
AESTHETIC			
AES1	The design of interface on the website is attractive and it is pleasing to look at screens.	0.695	7.113
AES2	I feel comfortable with the colors and graphical illustrations used on the web site (e.g. colors, pictures).	0.745	-
RESPONSIVENESS			
RES1	I can reach the help documents in case of necessity	0.781	13.76
RES2	This website provides clear and useful messages if I don't know how to proceed.	0.771	13.57
RES3	Web page always provides me to return to first or previous page at any time (e.g. links to homepage) in case of necessity.	0.765	-
RES4	Messages of error or warning are understandable and prevent possible errors from reoccurring.	0.806	14.52
RES5	There are security message and e-mail notification on the system.	0.775	14.27

–: Fixed for estimation.

*** All values are significant at the 0.01 level.

For convergent validity factors loading and AVE should be greater than 0.5.Also CR should be greater than 0.7. If we look at Table 4.9, all factor loadings and AVE

values are greater than 0.5 and all CR values expect for Aesthetics supporting convergent validity.

Moreover, we tested discriminant validity with 23 questions for both model. The square root of AVE in each diagonal has been marked (*) in Table 4.10 explaining that the square root of AVE for a construct is greater than the correlations involving the all constructs. In addition, AVE value for each construct is greater than 0.50 as seen AVE values in Table 4.9.

Table 4.9 Factor loadings, AVE, root-square of AVE and CR.

Dimensions	Items	Estimate (λ)	AVE	$\sqrt{(AVE)}$	CR
EASINESS	EASY1	0.719	0.544	0.738	0.781
	EASY2	0.778			
	EASY3	0.714			
CLARITY&VISIBILITY	CV1	0.817	0.602	0.776	0.858
	CV2	0.829			
	CV3	0.737			
	CV4	0.714			
RESPONSIVENESS	RES1	0.781	0.608	0.780	0.886
	RES2	0.771			
	RES3	0.765			
	RES4	0.806			
	RES5	0.775			
AESTHETICS	AES1	0.695	0.519	0.720	0.683
	AES2	0.745			
CONTROLLABILITY	CONT1	0.797	0.630	0.794	0.773
	CONT2	0.79			
COMPLETENESS	CMP1	0.714	0.578	0.760	0.804
	CMP2	0.795			
	CMP3	0.769			
INFORMATION QUALITY	IQ1	0.781	0.600	0.775	0.750
	IQ2	0.768			
EFFICIENCY	EF1	0.837	0.691	0.832	0.818
	EF2	0.826			

Table 4.10 Discriminant validity of dimensions.

Dimensions	Easiness	Clarity and visibility	Responsiveness	Aesthetics	Controllability	Completeness	Information Quality	Efficiency
Easiness	0.737*							
Clarity and Visibility	0.673	0.775*						
Responsiveness	0.394	0.393	0.779*					
Aesthetics	0.448	0.487	0.26	0.720*				
Controllability	0.158	0.085	0.543	-0.072	0.793*			
Completeness	0.66	0.58	0.551	0.303	0.306	0.760*		
Information Quality	0.688	0.466	0.496	0.372	0.213	0.688	0.774*	
Efficiency	0.296	0.372	0.66	0.176	0.494	0.623	0.393	0.831*

In addition, unidimensionality test for each factor has been checked in order to see whether or not they load onto one factor. As seen in Table 4.11, Eigenvalues are greater than 1 and Second Eigenvalues are smaller and far away from 1, which support the construct validity as well.

Table 4.11 Unidimensionality of the checklist factors.

Dimensions	Number of Items	First Eigenvalue	Second Eigenvalue
Easiness	3	2.083	0.503
Information Quality	2	1.6	0.4
Aesthetics	2	1.518	0.482
Clarity and visibility	4	2.789	0.501
Responsiveness	5	3.371	0.544
Completeness	3	2.147	0.488
Controllability	2	1.63	0.37
Efficiency	2	1.691	0.309

4.4.3 Analysis of Case Study

The findings in this article showed that QUIN methodology has a strong ability to find major and minor usability problems for informational web sites. The methodology has been constructed with respect to exploratory and confirmatory factor analysis. The instrument showed the construct validity with eight factors and 23 questions. Moreover, the results of the study also confirmed that QUIN has a good reliability.

Analysis of the usability evaluation brought a new perspective to light regarding the web site interface. The proposed methodology identified the most critical usability problems. Although usability experts would like to discover all usability problems, they would fix them in accordance to the level of severity for the users, considering budget and time limitations. According to QUIN methodology, Completeness proved to be the most critical dimension.

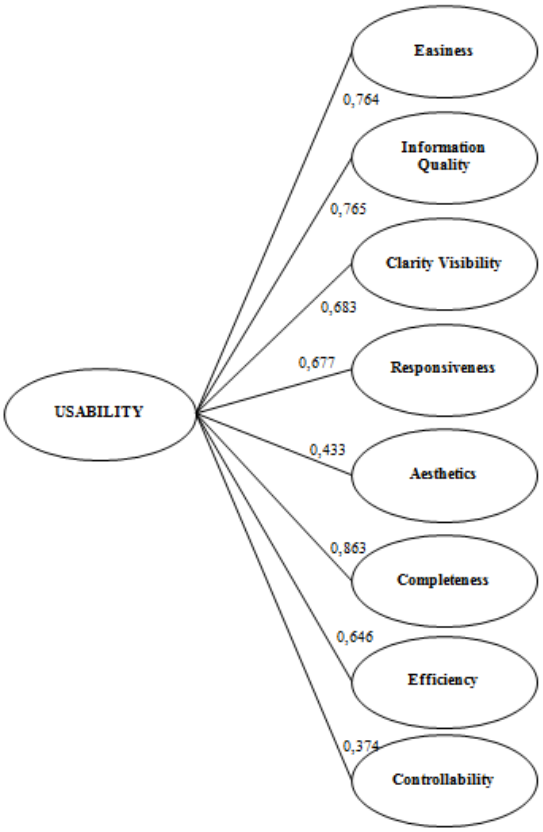


Figure 4.2 Regression weights for QUIN methodology.

When the regression weights of CFA results for Completeness are analyzed in Figure 4.2, the most important item evaluating this dimension is CMP2; i.e. “Pages,

titles, and links have been gathered in order(Clustering, grouping and sub-links)". The users get confused while they navigate through the pages because pages, titles, links are not grouped as it is supposed to be**.**

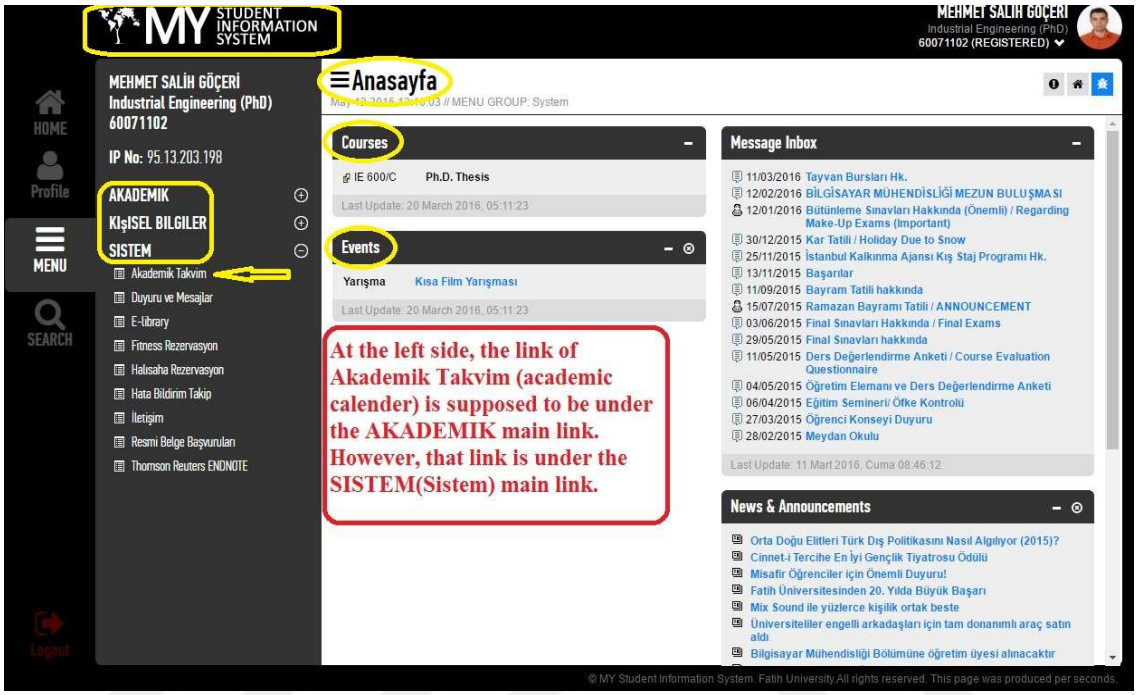


Figure 4.3 Default Interface of the student portal.

In Figure 4.3, even though language of the web portal was English as default, page title was "Anasayfa" which is in Turkish, some head titles are in English (see Courses, Events) and links on the left are in Turkish. When home page (Anasayfa) is clicked, some links on the left appear as Turkish: Akademik (Academic), KIşISEL BILGILER (Personal Information) and Sistem(System). Under "Sistem", there is a link "Akademik Takvim" (Academic Calendar) which should be under "Akademik" link which is much more appropriate place rather than "Sistem". In addition in Figure 4.3, the two links at the left side, namely "Kisisel Bilgiler" (under the KIşISEL BILGILER main link) and "Profile", both provide almost the same features regarding personal information. Moreover, in Figure 4.5, under "KIşISEL BILGILER" link, there is another link called "Kisisel Bilgiler". These situations are very confusing for the users because the same words are used for different links for the same purposes.

Second important item in Completeness dimension is CMP3 which is slightly different from CMP1: “The activity icons, buttons, labels, and links are designed with respect to the purpose intended.” This time, the language of the portal has been changed from English to Turkish as seen in Figure 4.4. There are many confusing words used for the same purpose on the web site. Therefore, users need to spend extra time on the website in order to achieve their goals such as registering for a class or requesting a document from student affairs. For instance, when users click the “Academic” link, they face three different registration links; “Course registration”, “REGISTRATION PROCESSES”, and “Registration Renewal”. Three different links regarding registration would mislead the users.

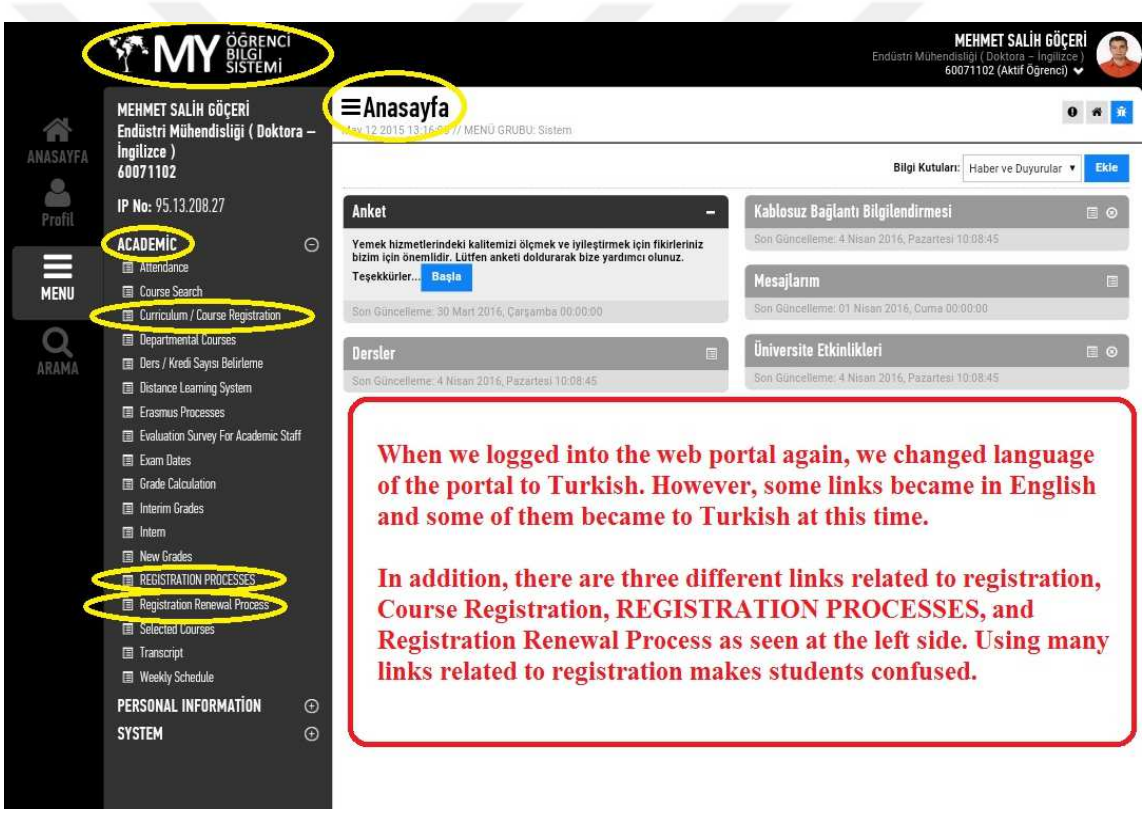


Figure 4.4 Interface of the student portal as Turkish

The third problem in Completeness is CMP1, “The website contains similar names, the same structure across the pages” which supports the previous findings. Users face problems while navigating through the web site because if the web portal does not contain similar names, the same structure across pages, users would get lost and spend extra time to find what they need. For example in Figure 4.5, the links on the left side

“Ayarlarım” (“my settings” in English) and “Profile” are exactly the same They both opens the password, email, and phone updates.

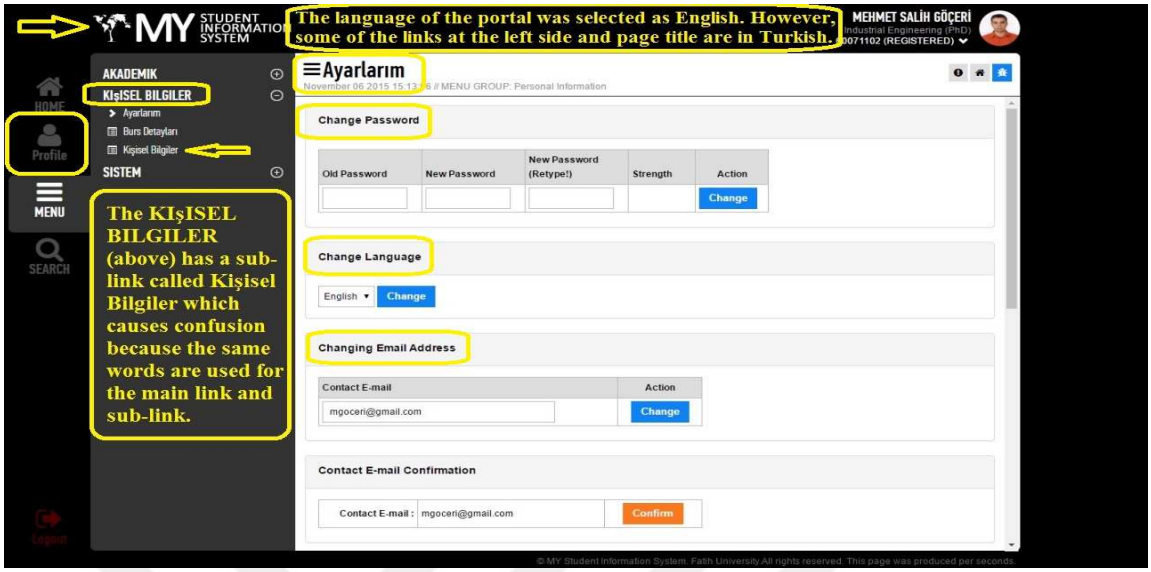


Figure 4.5 Interface of the student portal as English.

As mentioned above in Figure 4.5, if user clicks on “KIşISEL BILGILER” main link and then “Kişisel Bilgiler” sub-link, web portal will open personal settings such as changing email, phone, and password. Users can reach this part also by clicking “Profile”. Different name and terminology used across the page for the same purposes which may cause confusion.

In order to improve usability of the web portal, web designers should rather change the design of MyFatih portal considering the clustering of the web site links, avoiding the usage of the same phrases for different purposes for improving systems completeness.

On the other hand, the least important dimensions for the user interface have been found as “Controllability” (the least important) and “Aesthetics”. Controllability items were regarding redesigning the interface according to user preferences. The item with smallest regression weight in “Controllability” is “It is possible to increase and decrease the size of the font”. That means it does not seem to be a major issue for the users. The other question in this dimension is” The web site is personalized for each visitor.” This explains that users do not pay attention much to the customization of the student portal. If we look at the demographics of the participants, only 28% of the users visit the portal more than once a day.

All dimensions with their respective weights which determine the severity of usability problems are summarized below for the model:

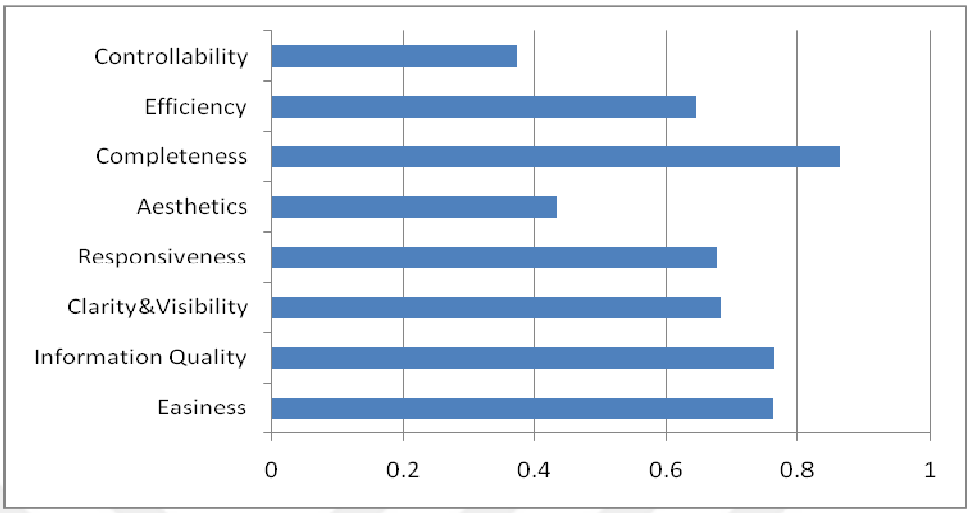


Figure 4.6 Dimensions and respective regression weights.

It is obvious that the most critical dimensions to be considered are Completeness, Easiness and Information Quality respectively. On the contrary, the least critical dimensions are Controllability, Aesthetics, and Efficiency respectively. Moreover, these finding support the idea of Oztekin et al, (2009) because they used a questionnaire with some experimental observations, in order to find the usability problems of the same user interface (the previous version of the student portal). However, they came up with the conclusion that Efficiency has less weight on usability, as found in QUIN methodology. The model proved that Efficiency is not highly critical for usability without performing lengthy experimental observation.

In addition, the second least important factor is found to be “Aesthetics”. This finding supports that why Controllability dimension was the least important factor for users because if they had control over the web portal by customizing it according to their own preferences (e.g. changing colors, screen resolution, font size etc.), they would make the portal more attractive (Aesthetics). The reason why these two dimensions were found to be the least important factors could be interpreted as the users’ indifference in controllability of the interface (?). More deliberate analysis is required to ensure this finding.

CHAPTER 5

COMPARISON OF FINDINGS

5.1 QUALITATIVE VERSUS QUANTITATIVE DATA

As a matter of fact, no single technique is capable of finding all usability problems. Therefore, user feedback has been checked as well to crosscheck the usability problems in another perspective. 326 participants' feedback from surveys has been collected as qualitative data. Some participants mentioned only a single problem while others listed more. 20 participants did not give any feedback on questionnaire. 441 problems have been mentioned by users. A number assigned to the each problem in order to categorize them. Problem groups have been named according to dimension names used in QUIN methodology. Since user feedback provided more problems, the name of problem groups were more than the number of dimensions in QUIN methodology.

A total of 14 different problem groups mentioned by the participants has been found. Some users used synonymous words or implied the same problems with different words and so on. For example, the issues for "Flexibility of the system" are the problems of printing, course selection, language selection, and password changes as seen in Table 5.1.

Table 5.1 Categories and frequencies of the problems according to user feedback.

Category of the Problems	Frequency of the Problem	Percentage
Search Option	78	18%
Completeness	69	16%
Aesthetics	52	12%
Efficiency (Speed)	51	12%
Clarity and Visibility	45	10%
Easiness	23	5%
Controllability	23	5%
Responsiveness and Help	18	4%
Mobile Phones Compability	18	4%
Information Quality	14	3%
Course selection	8	2%
Language Options	8	2%
Updateness	8	2%
Printing Issues	6	1%
Announcements	4	1%
System Surveys Frequency	6	1%
Password Changes Issues	3	1%
Security Concern	6	1%
Accessibility to Online Library	1	0%
Total	441	100%

These problems grouped under “Flexibility of the System” as seen in Table 5.2. Additionally, users mentioned “Mobile Phones Compatibility” issue because they express that student portal is not working in some smart devices. In addition, users mentioned that student portal should have a mobile application on the app stores so that they can download and use with their smart phones. Even though these two dimensions might be under the “Flexibility of the System”, we decided to split them in different groups because “Mobile Phones Compatibility” might be an issue of the smart phones rather than the web system and whether or not having an application in the app stores is an issue mobile phone rather than flexibility of the system.

After categorizing the problems into 14 different clusters, the number of problems related to the same cluster has been counted to find out the frequency of each problem group. Additionally, the relevant percentages of each category has been calculated as seen in Table 5.1

Table 5.2 Rearranged categories and frequency of the problems.

Category of the Problems	Frequency of the Problem	Percentage
Search Option	78	17.7%
Completeness	69	15.6%
Aesthetics	52	11.8%
Efficiency (Speed)	51	11.6%
Clarity and Visibility	45	10.2%
Flexibility of the System	25	5.7%
Easiness	24	5.4%
Controllability	23	5.2%
Information Quality	22	5.0%
Responsiveness and Help	18	4.1%
Mobile Phones Compatibility	18	4.1%
Security Concern	6	1.4%
System Surveys Frequency	6	1.4%
Announcements	4	0.9%
Total	441	100%

As mentioned above, since the search button in the website was not working, it was not a surprise that “Search Option” has been found as the most important usability problem. The second most important usability factor is “Completeness” as identified by quantitative data. In addition, if we add search option into Completeness dimension, since it could be regarding the web site completeness issue, 34% of the problems mentioned is found to be related to Completeness; supporting the QUIN survey results. However, the second important factor according to user feedback is “Aesthetics” while it is “Information Quality” in the survey quantitative results. In addition, third important factor mentioned in user feedback is “Efficiency” while it is “Easiness” in survey quantitative findings.

However, the QUIN quantitative result covers 67.1% of the total problems, which is a great finding. Additionally, if “Search” has been added into Completeness dimension, it will cover 87% of the total problems. In fact, search dimension were in QUIN questionnaire before starting the exploratory factor analysis. However, because of the fact that the search button in the student portal was not working, data obtained from “Search” dimension produced an inconsistent result for EFA. Therefore, “Search” dimension was deleted even though it was included in QUIN. Therefore, it can be

concluded that QUIN quantitative results have a capability to discover up to 87% of the usability problem which has been found by user feedback

Additionally, using qualitative data allowed us to detect the more usability problems from user perspective rather than survey's quantitative results. Moreover, Aesthetics and Efficiency dimensions are found to be more important in terms of qualitative data than quantitative results. Interestingly, "Flexibility of the System" and "Mobile Phones Compatibility" is found a quite common issue for users which did not exist as separate dimensions in the questionnaire. These two extra findings are an indicator showing how mobile platforms and the flexibility of the systems are becoming more important for students who require more flexible and usable systems for smart devices.

CHAPTER 6

CONCLUSION

Qualitative feedbacks from users strongly support the QUIN's quantitative results. Although data collection with a questionnaire could be easier, it cannot determine all usability problems. Therefore, a questionnaire including user feedbacks has the potential to identify majority of the usability problems. Using a questionnaire with user feedbacks was an enriched approach to uncover usability problems of the portal.

Additionally, developing a questionnaire well-supported from usability guidelines and measurement models could help to construct a well-design usability measurement model which could have a strong problem finding capability for the usability problems.

Therefore, before constructing a usability evaluation questionnaire, detailed-literature review process is a mandatory, however, it should be supported with another evaluator such as user feedback.

To sum up, majority of the usability evaluation research used questionnaires without qualitative data support; thus, they could not detect all kinds of usability problems. Therefore, user feedback should be included in the questionnaires.

The distinctive contributions of the proposed QUIN methodology can be listed as follows:

1. Using a questionnaire supported with user feedback discover the major and minor usability problems for web sites.

2. The proposed methodology enables to compare the performance of quantitative versus qualitative findings.

3. The QUIN methodology brings a new perspective to discover the most crucial usability problems with qualitative data support.

The major limitation of this study is that the proposed methodology has been applied to a single web portal. A comparative study with several student portals would provide a better validation for the QUIN methodology. Additionally, incorporating administrative and academic staff in the study would provide a deeper perspective for detecting of the usability problems. These concerns can be investigated as future research.

Acknowledgements

The authors acknowledge the school officials of Fatih University who helped us to reach sufficient number of participants to convey the survey and improve the quality of the study with their tremendous efforts, valuable comments, and suggestions.

REFERENCES

Andrews, K., *Human-Computer Interaction*, http://courses.iicm.tugraz.at/hci/hci.pdf 2016.

Abran, A., Khelifi, A., Suryn, W. and Seffah A, "Usability meanings and interpretations in ISO standards", *Software Quality Journal*, 2003.

Ariola, M.M., *Principles and Methods of Research*, 2006 Ed., Manila: Rex Book Store, 2006.

Bandalos, B., *Confirmatory factor analysis*, Applied multivariate statistics for the social sciences (3rd edition), pp. 389-420, Mahwah, NJ: LEA, 1996.

Becker, S.A. and Mottay, F.E., "A global perspective on web site usability", *IEEE Software 18 (1)*, pp. 54-61, Jan/Feb 2001.

Belanche D., Casalo, L.V. and Guinaliu, M., "Website usability, consumer satisfaction and the intention to use a website: The moderating effect of perceived risk", *Journal of Retailing and Consumer Services 19*, pp. 124–132, 2012.

Benbunan-Fich, R., "Using protocol analysis to evaluate the usability of a commercial web site", *Information Management*, Vol. 39, No.2, pp. 151-163, 2001.

Bennett, J.L., "The user interface in interactive systems", *Annual Review of Information Science 7*, pp. 159-196, 1972.

Bevan, N., Kirakowski, J. and Maissel, J., " What is usability?", *in: Proceedings of the Fourth International Conference on Human-Computer Interaction*, pp. 651-655, Stuttgart, Germany, 1991.

Bevan, N., "Measuring usability as quality of use", *Software Quality Journal 4*, pp.115-150, 1995.

Bevan, N. and Azuma, M., "Quality in use: Incorporating human factors into the software engineering lifecycle", *In: Proceedings of the Third IEEE International Software Engineering Standards Symposium and Forum ISESS'97*, pp. 169-179, 1997.

Bhattacharya, Y., "Employee Engagement as a Predictor of Seafarer Retention: A Study among Indian Officers", *The Asian Journal of Shipping and Logistics*, Vol. 31, Issue 2, pp. 295–318, June 2015.

Biasutti, M., "The student experience of a collaborative e-learning university module", *Computers and Education 57*, pp. 1865–1875, 2011.

Brooke, J., SUS: A quick and dirty usability scale. *Usability evaluation in industry*, In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & I. L. McClelland (Eds.), London 1996.

Taylor and Francis, http://www.usability.serco.com/trump/documents/Suschapt.

doc

Brown, W., Yen, P., Rojas, M. and Schnall, R., "Assessment of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology", *Journal of Biomedical Informatics*, Vol. 46, Issue 6, pp. 1080–1087, December 2013.

Bruun, A. and Stage, J., "New approaches to usability evaluation in software development : Barefoot and crowd sourcing", *The Journal of Systems and Software 105*, pp. 40–53, 2015.

Burns, A.C. and Bush, R.R., *Marketing Research*, Prentice-Hall: NJ, 1995.

Cao, Q. and Dowlatshahi, S., "The impact of alignment between virtual enterprise and information technology on business performance in an agile manufacturing environment", *Journal of Operations Management, 23*, pp. 531-550, 2005.

Carman, J. M., "Consumer perceptions of service quality: An assessment of the SERVQUAL dimensions", *Journal of Retailing, 66(1)*, pp. 33–55, 1990.

Cassino, R., Tucci, M., Vitiello, G. and Francese, R., “Empirical validation of an automatic usability evaluation method”, *Journal of Visual Languages and Computing*, Vol. 28, pp. 1-22, 2014.

Castilla, D., Palacios, A.G., Miralles, I., Lopez,J.B., Parra, E., Berges, S.R., Botella, C., “Effect of Web navigation style in elderly users”, *Computers in Human Behavior 55*, pp. 909-920, 2015.

Chau, M. and Wong, C.H., "Designing the user interface and functions of a search engine development tool", *Decision Support Systems 48 (2*), pp. 369-382, 2010.

Chen, S.Y. and Macredie R.D., "The assessment of usability of electronic shopping: A heuristic evaluation”, *International Journal of Information Management 25*, pp. 516–532, 2005.

Chiagouris, L. and Wansley, B., "Branding on the Internet", *Marketing Management*, pp. 34-38, July 2000.

Choin, J.H. and Lee, H.J., “Facets of simplicity for the smart phone interface: A structural model, *International Journal of Human-Computer Studies*, 2011.

Cronin, J.J. and Taylor, S.A., “Measuring Service Quality: A Reexamination and Extension”, *Journal of Marketing*, 56(3), p. 55, Jul 1992.

Conyer, M., “User and usability testing: How it should be undertaken”, *Australian Journal of Educational Technology*, 11(2), pp. 38-51, 1995.

Couper, M. P., "Web surveys: A review of issues and approaches", *Public Opinion Quarterly, 64*, pp. 464-494, 2000.

Cronbach, L., *Essentials of Psychological Testing*, Harper and Row, New York, 1970.

De Vaus, D., *Surveys in Social Research*, 5th edition, Social Research Today, 2001.

Delice E.K. and Gungor, Z., “The usability analysis with heuristic evaluation and analytic hierarchy process”, *International Journal of Industrial Ergonomics 39*, pp. 934–939,2009.

Downing, C. E. and Liu, C., "Assessing web site usability in retail electronic Commerce", *IEEE 35th annual computer software and applications conference (COMPSAC)*, pp. 144-151, 2011.

Dumas. J., "The great leap forward: The birth of the usability profession", *Journal of Usability Studies, 2(2)*, pp. 54-60, 2007.

Dumas, J.S. and Redish, J.C., *A Practical Guide to Usability Testing*, 1st Edition, Intellect Books Exeter, UK, 1999.

Elling, S., Lentz, L. and De Jong, M., "Retrospective think-aloud method: Using eye movements as an extra cue for participants' verbalizations", *Proceedings of the 2011 annual conference on human factors in computing systems : ACM*, pp. 1161–1170, New York, 2011.

Elling, S., Lentz, L., De Jong, M. and Bergh, H.V.D., "Measuring the quality of governmental websites in a controlled versus an online setting with the Website Evaluation Questionnaire", *Government Information Quarterly 29*, pp. 383–393, 2012.

Fabrigar, L.R., Wegener, D.T., MacCallum, R.C. and Strahan, E.J., "Evaluating the use of exploratory factor analysis in psychological research", *Psychological Methods, 4(3)*, pp. 272-299, 1999.

Fan, X., Thompson, B. and Wang, L., "Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes", *Structural Equation Modeling: A Multidisciplinary Journal 6 (1)*, pp. 56-83, 1999.

Fernandez, A., Insfran, E. and Abrahao, S., "Usability evaluation methods for the web: A systematic mapping study", *Information and Software Technology 53*, pp. 789–817, 2011.

Field, A., *Discovering Statistics using SPSS*, London: Sage, 2009.

Flavian, C., Guinaliu, M., and Gurrea, R., "The role played by perceived usability, satisfaction and consumer trust on website loyalty", *Information & Management, 43 (1)*, pp.1-14, 2006.

Fogli, D. and Guida, G., "A practical approach to the assessment of quality in use of corporate web sites", *The Journal of Systems and Software 99*, pp. 52-65, 2014.

Folmer, E. and Bosch, J., "Architecting for usability: a survey", *Journal of Systems and Software 70 (1-2)*, pp. 61-78, 2004.

Fornell, C. and Larcker, D.F., "Evaluating structural equation models with unobservable variables and measurement error", *Journal of Marketing Research 18 (1)*, pp.39-50, 1981.

Fu, L., Salvendy, G. and Turley, L., "Effectiveness of user testing and heuristic evaluation as a function of performance classification", *Behaviour and Information Technology 21 (2)*, pp. 137-143, 2002.

Galesic, M. and Bosjnak, M., "Effects of questionnaire length on participation and indicators of response quality in a web survey", *Public Opinion Quarterly 73*, pp. 349-360, 2009.

Garvin, D.A., *Managing Quality: The Strategic and Competitive Edge Hardcover*, Free Press, 1st Edition, February 1988.

George, D., and Mallery, P., *SPSS for Windows step by step: A simple guide and reference*, 11.0 update (4th ed.), Boston: Allyn & Bacon, 2003.

Gorsuch R.L., *Factor analysis*, (2nd Edition), Hillsdale, NJ: Erlbaum, 1983.

Grönroos, C., "A Service Quality Model and its Marketing Implications", *European Journal of Marketing 18 (4)*, pp. 36-44, 1984.

Gupta, A. and Chen, I., "Service quality: implications for management development", *International Journal of Quality & Reliability Management 12 (7)*, pp. 28-35, 1995.

Hair, F., Anderson, R., Tatham, R. and Black, W., *Multivariate Data Analysis with Readings*, 4th Edition, Prentice-Hall International, London, 1995.

Hair, J.F., Black, W.C., Babin, B.J. and Anderson, R.E., *Multivariate Data Analysis*, 7th Edition, Prentice Hall, Upper Saddle River, New Jersey, 2010.

Hallahan, K., "Improving public relations web sites through usability research", *Public Relations Review, 27(2),* pp. 223-239, 2001.

Harvey, C., Stanton, N.A., Pickering, C.A., McDonald, M. and Zheng, P., "In-Vehicle Information Systems to meet the needs of drivers", International Journal of Human-Computer Interaction 27, pp. 505-522, 2010.

Heerwegh, D. and Loosveldt, G., "Face-to-face web surveying in a high-internet coverage population, Differences in response quality", *Public Opinion Quarterly 72*, pp. 836-846, 2008.

Hill, R., "What sample size is "enough" in internet survey research", *Interpersonal Computing and Technology: An Electronic Journal for the 21st Century*, *6*, pp. 3-4, 1998.

Hooper, D., Coughlan, J. and Mullen, M., "Structural Equation Modelling: Guidelines for Determining Model Fit", *Electronic Journal of Business Research Methods, 6(1),* pp.53-60, 2008.

Hornbaek, K., "Current practice in measuring usability:Challenges to usability studies and research", *Kasper Hornbæk_Department Int. J. Human-Computer Studies 64,* pp.79-102, 2006.

Insfran, E. and Fernandez,A., "A Systematic Review of Usability Evaluation in Web Development1", *ISSI Group, Department of Information Systems and Computation*, pp. 81, s/n, 46022, Valencia, Spain, 2008.

Isaac, S. and Michael, W.B., *Handbook in research and evaluation*, San Diego, CA: Educational and Industrial Testing Services, 1995.

ISO 9241-10, Ergonomic requirements for office work with visual display terminals (VDT's), Part 10: Dialogue principles, 1996.

ISO 9241-11, Ergonomic requirements for office work with visual display terminals (VDTs)-Part 11: guidance on usability, 1998.

ISO/IEC 9126-1, Software engineering, Product quality, Part 1: Quality model, 2001.

ISO/IEC 25010, Systems and software engineering, Systems and software Quality Requirements and Evaluation (SQuaRE), System and software quality models, 2011.

Iwaarden, J., Wiele, T., Ball, L. and Millen, R., "Perceptions about the quality of web sites: a survey amongst students at Northeastern University and Erasmus University", Information & Management, Vol. 41, Issue 8, pp. 947–959, November 2003.

Julious, S.A., "Sample size of 12 per group rule of thumb for a pilot study", *Pharmaceutical Statistics 4*, pp. 287-291, 2005.

Jolliffe, I.T., "Discarding variables in a principal component analysis 1: artificial data", *Applied Statistics*, Vol.21, pp.160-173, 1972.

Jöreskog, K. and Sorbom, D., *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language*, IL: Scientific Software International Inc., Chicago, 1993.

Juristo, N., Moreno, A.M. and Sanchez-Segura, M.I., "Analyzing the impact of usability on software design", *Journal of Systems and Software 80 (9),* pp. 1506-1516, 2007.

Kaiser, H.F., "The application of electronic computers to factor analysis", *Educational and Psychological Measurement, 20*, pp.141-151, 1960.

Kaufman D., Roberts W.D., Merrill J., Lai T.Y. and Bakken S., "Applying an evaluation framework for health information system design, development, and implementation", *Nurs Res 2006, 55* (2 Suppl), pp. 37–42, 2006.

Keevil, B., “Measuring the usability index of your web site”, *In: Proceedings of the 16th Annual International Conference on Computer Documentation, SIGDOC e ACM Press*, pp. 271-277, 1998.

Kim, J. and Lee Y., “From design features to financial performance: a comprehensive model of design principles for online stock trading sites”, *Journal of Electronic Commerce Research 3(3)*, pp. 128–143, 2002.

Kirakowski, J., Claridge, N., and Whitehand, R., “Human-centered measures of success in website design”, *Proceedings of the 4th conference on human factors & the web*, pp. 1–9, Baskerville, NJ, 1998.

Krug, S., *Don't make me think: Common sense approach to web usability,* (2nd ed.), pp. 135-136, New Riders Press, Berkeley, Ca, 2006.

Ladhari, R., “Developing e-service quality scales: a literature review”, *Journal of Retailing and Consumer Services 17(6)*, pp. 464–477, 2010.

Lallemand, C., “Toward a closer integration of usability in software development: A study of usability inputs in a model-driven engineering process”, *Proceedings of the 3rd ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS ’11)*, pp. 299–302, 2011.

Lee, S. and Koubek, R. J., “The effects of usability and web design attributes on user preference for e-commerce web sites”, *Computers in Industry 61*, pp. 329–341, 2010.

Lee, Y. and Kozar, K.A., “Understanding of website usability: specifying and measuring constructs and their relationships”, *Decision Support Systems 52*, pp. 450-463, 2011.

Leuthold, S., Avila, J.A.B. and Opwis, K., “Beyond web content accessibility guidelines: Design of enhanced text user interfaces for blind internet users”, *International Journal of Human-Computer Studies*, Vol. 66, Issue 4, pp. 257-270, 2007.

Levi, D.M. and Conrad, F.G., “Usability Testing of World Wide Web Sites”, *Research Papers of Bureau of Labor Statistics*, 1998.

Lewis, R.C. and Booms, B., “The marketing aspects of service quality”, *American Marketing Association Proceeding*, pp. 99-104, Chicago, 1983.

Lewis, J.R. and Sauro, J., “The factor structure of the System Usability Scale”, Lecture Notes in Computer Science, 5619, pp. 94–103, 2009.

Li, Y.N., Tan, K.C., and Xıe, M., “Measuring web service Quality”, *Total Quality Management, 13(5)*, pp. 685–700, 2002.

Liang, T. and Lai, H., “Electronic store design and consumer choice: an empirical study”, *Information Management 39*, pp. 431–434, 2000.

Mack, Z., and Sharples, S., “The importance of usability in product choice: A mobile phone case study”, *Ergonomics, 52(12)*, pp. 1514–1528, 2009.

Manning, H., McCarthy, J.C. and Souza, R.K., “Why Most Web Sites Fail”, *Interactive Technology Series*, Vol. 3 (7), Forrester Research, September 1998.

Marsico, M.D. and Levialdi, S., “Evaluating web sites: exploiting user’s expectations”, *Int. J. Human-Computer Studies 60*, pp. 381–416, 2003.

Masood, M. and Musman, A., “The Usability and its Influence of an e-Learning System on Student Participation”, *Procedia - Social and Behavioral Sciences*, pp. 2325–2330, Athens, Greece, 2015.

McNamara, P., *Usability and eLearning, final coursework essay*, 2003. http://nte.unifr.ch/IMG/pdf/PMcNamaraUsabilityElearning.pdf.

Miller, R.B., *Human ease of use criteria and their tradeoffs*, IBM Technical Report TR 00.2185, IBM Corporation, Poughkeepsie, NY, 1971.

Molich, R. and Nielsen, J., “Improving a human-computer dialogue”, *Communications of the ACM 33(3)*, pp. 338-348, 1990.

Muyllea, S., Moenaertb, R. and Despontinc, M., “The conceptualization and empirical validation of web site user satisfaction”, *Information & Management 41*, pp. 543–560, 2003.

Nam, J., “Understanding the motivation to use web portals”, *Computers in Human Behavior 36*, pp. 263–273, 2014.

Navimipour, J. and Milani, S., “A comprehensive study of the resource discovery techniques in peer-to-peer net works”, *Networking and Applications*, pp. 1–19, 2015.

Navimipour, N.J. and Zareie, B., “A model for assessing the impact of e-learning systems on employees’ satisfaction”, *Computers in Human Behavior 53*, pp. 475–485, 2015.

Newman, W. And Taylor, A., “Towards a methodology employing critical parameters to deliver performance improvements in interactive systems”, *In: Proceedings of IFIP TC.13 International Conference on Human-Computer Interaction.* IOS Press, pp. 605-612, Amsterdam, 1999.

Nielsen, J., *Usability Engineering*, Academic Press, San Diego, CA, 1993.

Nielsen, J., *Usability inspection methods,* In*:* Proceedings of the CHI.ACM Press, 1994.

Nielsen, J., *Web Usability*, Edizione Apogeo, Milano, 2000.

Nielsen, J., “Usability 101: Introduction to Usability”, Nielson Norman Group, 04 January 2012.

Nielsen, J. and Levy, J., “Measuring usability: preference vs. Performance”, *Communications of the ACM 37(4)*, pp. 66-75, 1994.

Ozok, A.A., “Survey Design and Implementation in Human Computer Interaction”, *The Human-Computer Interaction Handbook,* 2nd edition, pp. 1151-1169, Lawrence Erlbaum Associates, Mahwah, 2008.

Oztekin A., Nikov A. and Zaim S., “UWIS: An assessment methodology for usability of web-based information systems”, *Journal of Systems and Software 82 (12)*, pp. 2038-2050, 2009.

Oztekin A., Delen D., Turkyilmaz A. and Zaim S., “A machine learning-based usability evaluation method for eLearning systems”, *Journal of Systems and Software 82 (12)*, pp. 2038-2050, 2013.

Parasuraman, A., Zeithaml, V. and Berry, L.L., “A conceptual model of service quality and its implications for future research”, *Journal of Marketing*, autumn, Vol. 49, pp. 41-50, 1985.

Parasuraman, A., Zeithmal, V. A. and Berry, L.L., “SERVQUAL: A multi-item scale for measuring consumer perceptions of service quality”, *Journal of Retailing, 64(1)*, pp. 12–37, 1988.

Quesenbery, W., *Choosing the right usability technique: Getting the answers you need*, http://www.wqusability.com/handouts/righttechnique-uf2008.pdf, 2008.

Ratner, J., *Human Factor and Web Development*, 2nd Edition, pp. 66, 2003.

Regan, W.J., “The Service Revolution”, *Journal of Marketing*, 27(3), pp. 57-62, July 1963.

Rigdon, E.E., “Structural equation modeling. In: Marcoulides, G. (Ed.)”, *Modern Methods for Business Research*, Lawrence Erlbaum Associates, Mahwah, NJ, pp.251–294, 1998.

Roussel, N., *Looking back: a very brief history of HCI*, January 2014. http://interaction.lille.inria.fr/~roussel/publications/2014-looking-back.pdf

Rubin, J. and Chisnell, D., *Handbook of Usability Testing*, 2nd Edition, Wiley, 2008.

Saraph, J.V., Benson, P.G. and Schroeder, R.G., “An Instrument For Measuring The Critical Factors Of Quality”, *Decision Sciences (Fall) 20(4)*, pp. 810-829, 1989.

Satorra, A. and Bentler, P.M., “Corrections to test statistics and standard errors in covariance structure analysis”, *Latent Variables Analysis: Applications for Developmental Research*, CA: Sage, 1994.

Schumacker, R.E. and Lomax, R.G., *A beginner's guide to structural equation modeling*, Mahwah, NJ: Lawrence Erlbaum Associates, 1996.

Seffah, A., Mohamed, T., Habieb-Mammar, H. and Abran, A., “Reconciling usability and interactive system architecture using patterns”, *Journal of Systems and Software 8 (11)*, pp.1845-1852, 2008.

Sengel, E., “Usability level of a university web site”, *Procedia - Social and Behavioral Sciences*, 4th International Conference on New Horizons in Education, Bursa, Turkey ,2013.

Shackel, B., *Ergonomics for a computer*, Design 120, pp. 36-39, 1959.

Shackel, B., “Usability-context, framework, definition, design and evaluation” In: Shackel, B., Richardson, S. (Eds.), *Human Factors for Informatics Usability*, Cambridge University Press, Cambridge, pp. 24, 1991.

Sharma, S., Richard M.D. and Oded G., "Identification and Analysis of Moderator Variables," *Journal of Marketing Research 18*, pp. 291-300, August 1981.

Sharma, S., *Applied multivariate techniques,* p.493, John Wiley and Sons, inc., New York, 1996.

Shneiderman B., *Designing the User Interface: Strategies for Effective Human-computer Interaction*, Addison-Wesley, MA, 1998.

Smith, S.L. and Mosier, J.N., *Guidelines for designing user interface software*, Mitre Report ESD-TR-86-278, The MITRE Cooperation, Bedford, MA, 1986.

Spool, J., De Angelo, T., Scanlon, T., Schroeder, W. and Snyder, C., *Web Site Usability*, Morgan Kaufmann, ISBN: 155860569X, 1999.

Spyridakis, J. H., Wei, C., Barrick, J., Cuddihy, E. and Maust, B., "Internet-based research: Providing a foundation for web-design guidelines", *IEEE Transactions on Professional Communication, 48*, pp. 242-260, 2005.

Teas, R.K., "Expectations, performance evaluation and consumer's perception of quality", *Journal of Marketing, 57(4)*, pp. 18–34, 1993.

Tedeschi, B., *Good Web Site Design Can Lead to Healthy Sales*, 1999. http://www.nytimes.com/ library/tech/99/08/cyber/commerce/30commerce.html.

Tourangeau, R., Rips, L.J. and Rasinski, K.A., *The psychology of survey response*, Cambridge: Cambridge University Press, 2000.

Tourangeau, R., "Cognitive aspects of survey measurement and mismeasurement", *International Journal of Public Opinion Research 15*, pp. 3-7, 2003.

Treiblmaier, H. and Pinterits, A., "Developing metrics for web sites", *The Journal of Computer Information Systems 50 (3)*, pp. 1-10, 2010.

Tullis, S.T. and Stetson, J.N., "A comparison of questionnaires for assessing Web site usability", *Usability Professionals Association (UPA) Conference*, pp. 7-11, Minneapolis, USA, 2004.

Tung, L.L., Xu, Y. and Tan, F.B., "Attributes of web site usability: a study of web users with the repertory grid technique", *International Journal of Electronic Commerce 13 (4*), pp. 97-126, 2009.

Turan, F. and Bayram O., "Information Access and Digital Library Use In University Students' Education: The case of Ankara University", *The 2nd International Conference on Integrated Information Procedia - Social and Behavioral Sciences 73*, pp 736 – 743, Ankara, Turkey , 2013.

Virzi, R.A., Sorce, J.F. and Herbert, L.B.A, "Comparison of Three Usability Evaluation Methods: Heuristic, Think-Aloud, and Performance Testing", *Proceedings of the Human Factors and Ergonomics Society*, 37th Annual Meeting, pp. 309-313, Seattle, WA, 1993.

Wagner, N., Hassanein, K. and Head, M., "The impact of age on website usability", *Computers in Human Behavior 37*, pp. 270–282, 2014.

Walji, M.F., Kalenderianb, E., Piotrowskia, M., Trana, D., Kookala, K.K., Tokedeb, O., Whitec, J.M., Vaderhoblic, R., Ramonid, R., Starke, P.C., Kimmesf, N.S., Lagerweijg, M. and Patelh, V.L. "Are three methods better than one? A comparative assessment of usability evaluation methods in an", *International journal of medical informatics 83*, pp. 361–367. 2014.

Westcott, R.T., *The Certified Manager of Quality/Organizational Excellence Handbook Quality Management Division*, 4th Edition, ASQ Quality Press, Milwaukee, Wisconsin, 2013.

Westbrook, K.W. and Peterson, R.M., "Business-to-business selling determinants of quality", *Industrial Marketing Management 27*, pp. 51-62, 1998.

Whitehead, C., "Evaluating web page and web site usability", *Proceedings of the 44th Annual Southeast Regional Conference*, pp. 788–789, 2006.

Wu, M., Wang, L., Li, M. and Long, H., "An approach of product usability evaluation based on Web mining in feature fatigue analysis", *Computers & Industrial Engineering 75*, pp. 230-238, 2014.

Yamane, T., "Statistics, an Introductory Analysis", 2nd Ed., New York: Harper and Row, 1967.

Yussof R.L., Paris T.N.S.T., Abas, H. and Zamand H.B., "Educational software especially in supporting the needs of learning for disabled students Mixed Usability Evaluation during the Development Cycle of MEL", *Procedia-Social and Behavioral Sciences*, Vol. 105, pp. 162-170, 2013.

Zeithaml V.A., "How Consumer Evaluation Processes Differ between Goods and Services", *Services Marketing*, 2nd Edition, Upper Saddle River, New Jersey: Prentice Hall, 1981.

Zeithaml, V.A., Parasuraman, A. and Berry, L.L., *Delivering Quality Service*, The Free Press, New York, 1990.

Zeithaml, V.A., Berry, L.L. and Parasuraman, A., "The Behavioral Consequences of Service Quality," *Journal of Marketing*, pp. 31-46, April 1996.

APPENDIX A

DECLARATION STATEMENT FOR THE ORIGINALITY OF THE THESIS, FURTHER STUDIES AND PUBLICATIONS FROM THESIS WORK

A.1 DECLARATION STATEMENT FOR THE ORIGINALITY OF THE THESIS

I hereby declare that this thesis comprises my original work. No material in this thesis has been previously published and written by another person, except where due reference is made in the text of the thesis. I further declare that this thesis contains no material which has been submitted for a degree or diploma or other qualifications at any other university.

Signature:
Date: June 2, 2016

A.2 FURTHER STUDIES

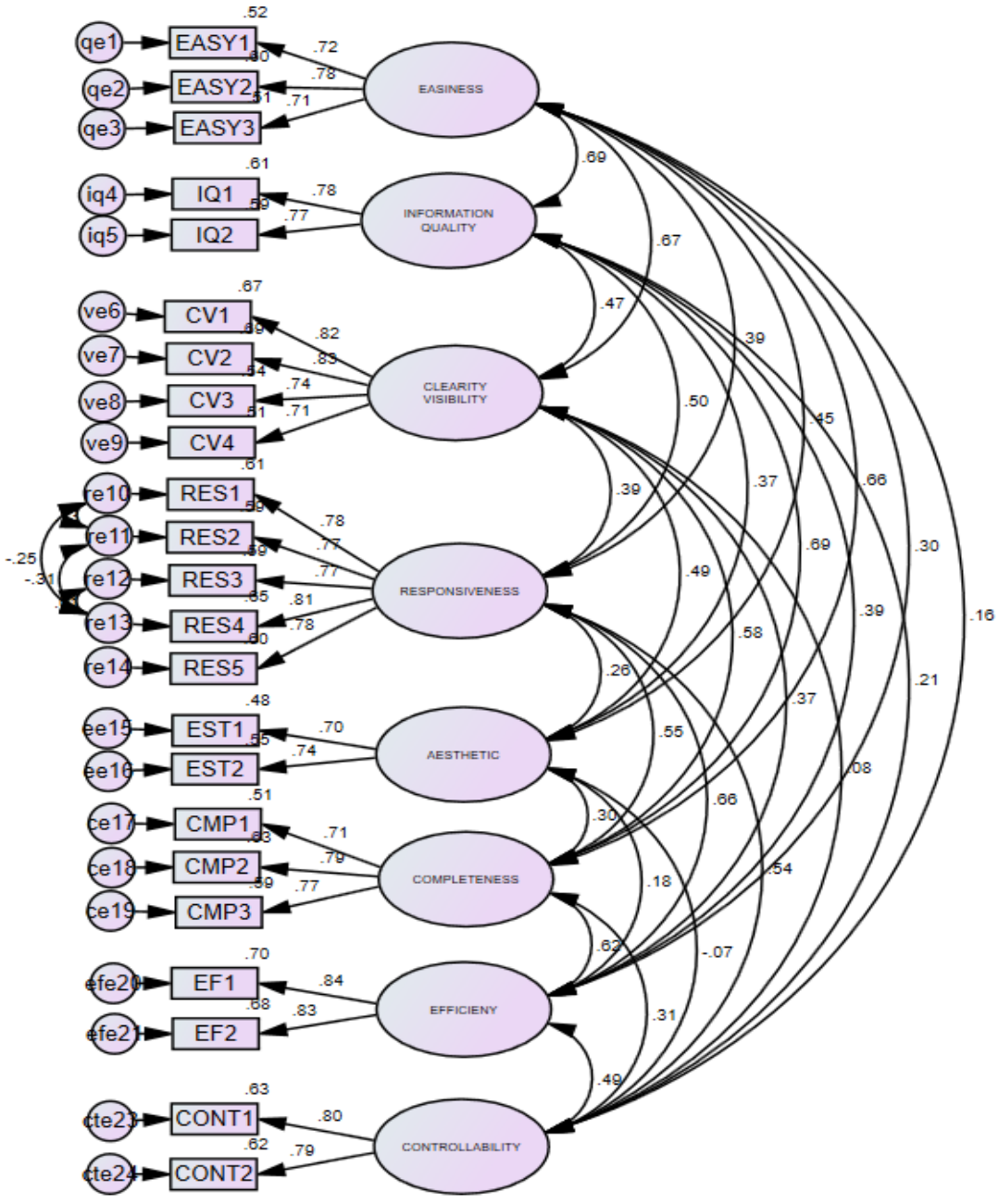
Further studies related to this thesis work can be done in the following area:

a. Using QUIN methodology for usability evaluation of user interface for faculty portal (PBS).

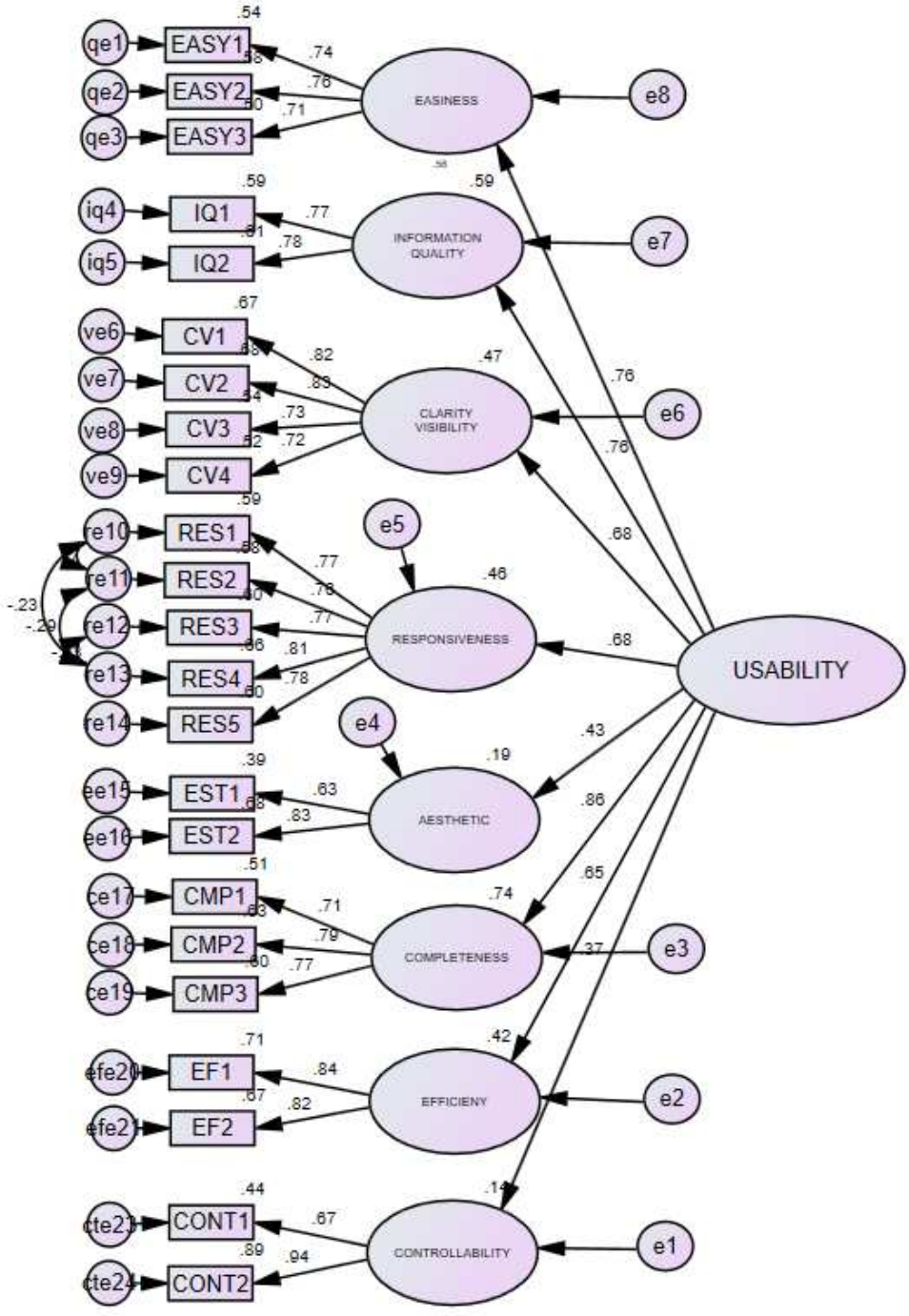
APPENDIX B

AMOS RESULTS FOR THE MODEL

B.1 AMOS STANDARDIZED WEIGHT FOR THE MODEL



B.2 AMOS RESULTS FOR SECOND ORDER LATENT FACTOR



CURRICULUM VITAE

CONTACT INFORMATION
Mehmet Salih GÖÇERİ
34500 Büyükçekmece, Istanbul, Turkey
Phone: (0212) 866 3300
Email: mgoceri@gmail.com

EDUCATION

Ph.D., Industrial Engineering, Fatih University, Istanbul, Turkey, June 2016
Dissertation: "A Usability Evaluation Methodology with a Questionnaire Supported with Qualitative Data for Web Site User Interface"

MBA, Marketing, Southern States University, Newport Beach, CA, USA, 2010

M.S., Industrial Engineering, Fatih University, Istanbul, Turkey, July 2008
Thesis: "Application of Supply Chain Principals in Optical Industry"

B.S., Industrial Engineering, Istanbul University, Istanbul, Turkey, July 2005

PROFESSIONAL EXPERIENCE

- Instructor, Leavey School of Business, MBA Department, Santa Clara University, Santa Clara, CA, USA (2016)
- Lecturer, Industrial Engineering Department, Fatih University, Istanbul, Turkey, (2015 - 2016)
- TSOL, Green Charter School, Greenville, South Carolina, USA, (February, 2015-November, 2015)
- Lecturer, Foreign Trade Department, Fatih University, Istanbul, Turkey, (2014 – 2014)

AWARDS AND HONORS

- Received full university scholarship for graduate studies in Industrial Engineering at Fatih University, (2012 - 2016)
- Received full university scholarship for graduate studies in Industrial Engineering at Fatih University, (2006 - 2008)

PUBLICATIONS

Conference Proceedings

Goceri, M.S., Gurcan, A, Ogcu G., Demirel, O.F., "Developing Industrial Engineering Ph.D. Program with AHP Based QFD", İzmir, Sep. 2012, 12.ci Uretım Arastırmaları Sempozyumu,