THE REPUBLIC OF TURKEY
BAHÇEŞEHİR UNIVERSITY

USABILITY METRICS ON E-LEARNING SYSTEMS

Master's Thesis

MEHMET RAFİ DOĞ

İSTANBUL, 2012

THE REPUBLIC OF TURKEY
BAHÇEŞEHİR UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
COMPUTER ENGINEERING

USABILITY METRICS ON E-LEARNING SYSTEMS

Master's Thesis

MEHMET RAFİ DOĞ

Supervisor: ASSOC. PROF.DR. ADEM KARAHOCA

İSTANBUL, 2012

THE REPUBLIC OF TURKEY
BAHÇEŞEHİR UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
COMPUTER ENGINEERING

Name of the thesis: USABILITY METRICS ON E-LEARNING SYSTEMS

Name/Last Name of the Student: Mehmet Rafi DOĞ

Date of the Defense of Thesis:

The thesis has been approved by the Institute of Graduate School in Sciences.

Assoc. Prof. F. Tunç BOZBURA
Graduate School Director
Signature

I certify that this thesis meets all the requirements as a thesis for the degree of Master of Science.

Assoc. Prof. Çağrı GÜNGÖR
Program Coordinator
Signature

This is to certify that we have read this thesis and that we find it fully adequate in scope, quality and content, as a thesis for the degree of Master of Science.

Examining Comitee Members	Signature
Thesis Supervisor Assoc. Prof. Adem KARAHOCA	-----------------------------------
Member Assoc. Prof. Alper TUNGA	-----------------------------------
Member Assoc. Prof. Yalçın ÇEKİÇ	-----------------------------------

ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisor Assoc. Prof. Dr. Adem Karahoca who helped me to discover many concepts that is used throughout the thesis. He was always very helpful and available when I needed help or advice, without his guidance and support, it would not be possible to complete this study. His help is greatly appreciated and I feel very lucky that I had the opportunity to work with him.

Also, I want to thank Dilek Karahoca who helped me a lot during this journey with sharing her ideas to improve my thesis and motivated with the belief of me.

Finally, I would like to express my gratitude to my mother (İnci Doğ), father (Yusuf Doğ) and Efsun Cebeci. Their love and support were essential for the completion of my thesis. It is to them I would like to dedicate this thesis.

ABSTRACT

USABILITY METRICS ON E-LEARNING SYSTEMS

Mehmet Rafi DOĞ

Computer Engineering Graduate Program

Supervisor: Assoc. Prof. Dr. Adem Karahoca

May 2012, 80 pages

Refined computer applications support learning in the cognitive domains and they are slightly different from commercial transaction processing applications. By the nature of such applications for evaluating their usability involves different methods.

In this thesis, we present the application and development of the framework for usability testing of interactive e-learning applications proposed by Masemola & de Villiers et al. (2006). In the first usability testing application, investigation the effectiveness of the think-aloud method is discussed and joined with co-discovery testing.

E-learning is evolving field, helping with the potential for increased reach and decreased cost of e-learning applications. Usability has crucial importance for e-learning due to the remote nature of the e-learning activities. In the direction to evaluate usability of e-learning, the applied usability factors are as follows: e-learning system feedback, consistency, error prevention, performance or efficiency, user like or dislike, error recovery, cognitive load, internationalization, privacy and online help.

To express the overall e learning usability, representation of all of the listed factors above is important and should be first discussed separately then combined into a single figure of merit.
We should take into account that all the dimensions listed are relatively imprecise and cannot be represented very easily such as numeric representation. Therefore, it is determined to use Fuzzy System Theory in order to represent them and Takagi-Sugeno (T-S) model is proposed by Takagi et al. (1985) and is used for Fuzzy reasoning by Wong et al. (2002) which provides single e-learning usability figure of merit while combining all the factors.

The components involved in this approach can be listed as: identifying the usability issues in e-learning context and testing for the Fuzzy logic based usability approach, combine the test results and change them into a single usability measure then validating the generated Fuzzy Model.

This thesis considers the use of fuzzy systems to illustrate each of these usability factors and shows how each of them affects the overall rate of the e-learning usability.

Key words: E-learning, Usability Metrics, Fuzzy Systems

ÖZET

UZAKTAN EĞİTİM SİSTEMLERİNDE KULLANILABİLİRLİK ÖLÇÜTLERİ

Mehmet Rafi DOĞ

Bilgisayar Mühendisliği Yüksek Lisans Programı

Tez Danışmanı: Doç. Dr. Adem Karahoca

Mayıs 2012, 80 Sayfa

Rafine bilgisayar uygulamaları bilişsel alanlarda öğrenmeyi desteklemektedir ve ticari işlem uygulamalarından biraz farklılık göstermektedir. Bu tür uygulamaların doğaları gereği kullanılabilirliklerini değerlendirebilmek için çeşitli yöntemler kullanılmaktadır.

Bu tezde, Masemola & de Villiers ve arkadaşları tarafından önerilen (2006) etkileşimli e-öğrenme uygulamalarının kullanılabilirlik testleri için uygulama ve geliştirme taslağı sunuyoruz. İlk kullanılabilirlik testi uygulamasında, soruşturma sesli düşünme yönteminin etkinliği tartışılmakta ve ortak bulma testi ile birleştirilmektedir.

E-öğrenme, artan erişim için potansiyele yardımcı ve e-öğrenme uygulamalarının maliyetini düşüren gelişmekte olan bir alandır.. Kullanılabilirlik e-öğrenme faaliyetlerinin uzak yapısı nedeniyle, e-öğrenme için hayati öneme sahiptir. E-öğrenmede kullanılabilirliğin değerlendirilmesi doğrultusunda, uygulanan kullanılabilirlik faktörler şunlardır: e-öğrenme sisteminde geribildirim, tutarlılık, hata önleme, performans ve verimlilik, kullanıcı beğenisi veya beğenmeme, hata kurtarma, bilişsel yük, uluslararasılaşma, gizlilik ve online yardım.

Genel olarak e öğrenme kullanılabilirliğini ifade etmek için, yukarıda sıralanan faktörlerin hepsinin nasıl temsil edildiği önemlidir ve ilk olarak ayrı ayrı belirtilmeli sonra da tek bir değer katsayısı altında birleştirilmelidir. Listelenen tüm boyutların göreceli olarak hatalı ve gösterimlerinin çok kolay bir şekilde örneğin sayısal gösterim olarak temsil edilemeyeceği hesaba katılmalıdır. Bu nedenle, bunları gösterebilmek için Bulanık Sistem Teorisi'nin kullanılmasına karar verilmiştir ve Takagi-Sugeno (TS) model Takagi ve arkadaşları tarafından önerilmiş (1985) ve Wong ve arkadaşları tarafından (2002) Bulanık mantık için kullanılmıştır, bu çalışma tek bir e-öğrenme kullanılabilirliği açısından bütün faktörleri birleştirerek tek bir değer katsayısı elde etmiştir.

Bu yaklaşımdaki bileşenler şu şekilde sıralanabilir: Kullanılabilirlik sorunlarını e-öğrenme açısından belirlemek ve bulanık mantık tabanlı kullanılabilirlik yaklaşımını test etmek, test sonuçlarını birleştirmek ve tek bir kullanılabilirlik ölçütüne çevirmek daha sonra da oluşturulan Bulanık Modeli doğrulamak.

Bu tez, bulanık sistemlerin kullanımını kullanılabilirlik faktörlerinin gösterimi için incelemekte ve her bir faktörün e- öğrenme kullanılabilirliğinin genel hızını nasıl etkilediğini göstermektedir.

Anahtar kelimeler: E-öğrenme, Kullanılabilirlik Ölçütleri, Bulanık Sistemler

CONTENTS

TABLES

FIGURES

ABBREVIATIONS

HCI	:	Human Computer Interaction
UEM	:	Usability Evaluation Method
HE	:	Heuristic Evaluation
MAS	:	Multi-agent Systems
UCD	:	User-centered Design
LCD	:	Learner-centered Design
AR	:	Action Research
MUSIC	:	Metric for Usability Standards in Computing
TLX	:	Task Load Index
WBL	:	Work-based Learning
UT	:	User Testing

1. INTRODUCTION

According to Teo et al. (2006) e-learning, by virtue of its unique distributed, asynchronous nature, shows much promise for fostering significant improvements in accessibility and opportunity to learn. It couples advances in technology and the advent of the information highway to eliminate barriers of time, distance, and socioeconomic status, thereby creating a whole new dimension of learning.

Because e-learning can transmit knowledge fast and effectively, many people accept it as a means of upgrading themselves and keeping up with the rapid changes that define the Internet. However, Lytras et al. (2002) and Polsani (2003) with all of e-learning's promise and the research effort spent on it, a literature review has unveiled major concerns about its effectiveness and appropriateness.

E-learning systems, as a powerful tool in modern education Doroodchi et al. (2008), can be evaluated from usability standpoint. Since usability of a tool determines how easy a tool can be used by its typical users, E-Learning usability was mainly studied from the interface perspective, whereas, from both the learner and instructor's perspectives the content of the course is more important.

From instructor's perspective, the major challenge in E-learning is the content development and corresponding assessment tools since the learning environment is very different from the traditional classroom. On the other hand, the learner has to deal with the content and instruction with an "offline" instructor as opposed to a live instructor. Consequently, Fetaji et al. (2007) proposed the quality of the content delivered digitally is a major factor for long term success of any E-learning endeavor.

International Organization for Standardization (ISO), (1998) has defined usability as the measurement of a product that can be used by particular users in order to achieve specific objectives with effectiveness, efficiency and satisfaction in a particular context. Most of the past researches on the usability of E-learning systems have been done in exploring the usability of interface of E-learning systems and also the links between usability characteristics and the E-learning endeavor success. However, research with respect to the usability factors within the context of E-learning content has not been given much attention.

The usability of computer software is defined by Shackel, (1981) measured by how easily and how effectively it can be used by a particular set of users, given particular kinds of support, to carry out a defined set of tasks, in a defined set of environments.

Ssembugabi et al. (2007) stated usability evaluation in the context of e-learning practices is a growing field, covers usability, interfaces and connecting with human-computer interaction (HCI), learning and pedagogy for education. Choice of usability evaluation methods (UEMs) are strongly connected and effected by time, effectiveness, efficiency, cost and implementation ease. In heuristic evaluation (HE) evaluation is determined by experts who are specialized in the domain area or human computer interaction. The reason of conducting evaluation study is investigating how to identify the usability problems in e-learning platform and compare the results with other evaluations via end-users.

Critical rate is identified on a consolidated usability problem set and applied while comparing the findings in the major and minor usability problems. The results show that the HE and survey results are in the same direction. In the findings phase four experts evaluated more problems than the remaining 61 users and identified 91 percent of the user's problems under the major problems consideration. It is proven by the experts that HE is a suitable, efficient and effective UEM for web-based learning implementations.

We introduce in this thesis how usability metrics applied to e-learning systems can lead to the definition of high adaptive learning/teaching systems. These systems will be able to adjust to a specific student and even they will be able to recommend to teachers how to improve the course.

Finally, to achieve our goal to create a highly adaptive teaching environment it will take some artificial intelligent techniques that will be modeled by means of multi-agent systems (MAS). An E-learning System is proposed, in which usability metrics are used to achieve intelligent behavior of the system to improve learning.

E-learning could be defined as learning activities involving computer networks. The networks could be intranets, extranets, or the Internet. E-learning has been made possible by the advancement of the Internet and multimedia technologies. However, it is the increasing number of learners, coupled with the limitation of resources available for education and training that has contributed to the growth of e-learning.

Ruttenbur et al. (2000) stated that the market for e-learning is forecasted to grow from $14.9 billion in 1998 to an estimated $33.7 billion in 2004. The increasing popularity of e-learning could be credited to a number of its important benefits. First, e-learning offers the potential for substantial cost savings in the provision of learning programs. Once developed, learning components such as learning materials and learning objects could be reused. Second, e-learning allows the learners the flexibility of learning anytime, anywhere instead of being restricted to a certain time and place. Finally, organizations can use e-learning to provide "just-in time" learning facilities where employees can access the right knowledge used by Ruttenbur et al. (2000) and Rossett et al. (2002) when they need it.

Due to the remote nature of e-learning, usability plays a vital role in the success of an e-learning program. An e-learning system with poor usability hinders e-learning efforts as the learners would spend more time on learning how to use the software instead of learning the needed contents. Yet, it is a concern raised by a growing number of experts Charles et al. (2003), Kapp, (2003) and Feldstein, (2002) in the current literature that usability and its evaluation have not yet been given due attention in the development of e-learning systems.

There is a real need for a usability evaluation method that is practical and specific to e-learning. Usability of a system "is measured by how easily and effectively it can be used by a specific set of users, given particular kind of support, to carry out a defined set of tasks, in a defined set of environments". The current approaches to evaluating the usability of a system could be categorized into: (a) Empirical testing; (b) Inspection; (c) Comparative usability measures; (d) Formal complex based measures. Empirical testing entails conducting usability evaluation in a fairly controlled situation to identify the problems that the user have with the system.

Inspection involves examining the design of various screens to identify problems that might occur in the user interfaces. Comparative usability measures consists of usability metrics that are designed to evaluate aspects of the interface such as the ability of the user to perform the tasks that s/he is required to do efficiently, and acceptance by the user of the user interface. Finally, formal complexity based measures on the other hand use a variety of criteria that rely on some formal or mathematical analysis of the user interface.

In this research we discuss the development of a comparative usability measurement approach for e-learning systems that has three components, namely, (a) heuristic evaluation, (b) testing aspect, (c) usability measure that consists of the results of the tests. The framework we employ for developing the usability measure is Fuzzy Systems

2. BACKGROUND

2.1 DEFINITION OF E-LEARNING

There exist several definitions for e-learning. In one hand, some of the definitions state a narrow view of the scope of e-learning, associating it to Internet and Web-based implementations; on the other hand Masemola et al. (2006) are stating broader definitions. As cited in de Villiers, (2005), Cedefop breaks down the e-learning in terms of methodology and formatting. For broad knowledge about definitions consists of electronic learning such as Web-based and CD-based approaches.

With this work, we consider the e-learning from the broader way, distinguish it as dealing with different forms of applications for instance, collaborative education implementations, educating games, open-ended learning platforms, hypermedia, learning management systems, simulations, web-based learning and the usage of tools for supporting learning.

According to Urdan et al. (2000) it is indicated that e-learning or electronic learning is the delivery of educational content via various electronic media, including the internet, intranets, extranets, satellite broadcast, audio and video tape, interactive TV, CD-Rom, interactive CDs, and computer-based training. E-learning is any learning, training or education that is facilitated by the use of well-known and proven computer technologies, specifically networks based on Internet technology.

E-learning system is an approach to facilitate and enhance learning activity by using computer and communications technology. Some of the methods used in e-learning are lectures, tutorials, performance measurement system and games. However, an effective e-learning always blend the methods together. Moreover, E-learning is useful for distance learning. Most of the education institution nowadays offers variety of courses that can be learned through modules provide in e-learning.

2.2 USABILITY EVALUATION IN E-LEARNING

The main reason for developing educational software is encouraging learning. From the users point of view, this developed software should be composed of different ways of learning methods by providing better usability for its users interacting with the system naturally and as perceptive as possible.

We can describe this interaction between learning system and users as a motivation for learning period and creates a synergy in between. If we take into account usability features, they should be appropriate for expected learning task, and they should not only permit users operating on the interactive system. After creation of the synergy between users and software, this software can be interpreted as integrated and establishing seamless connection the usage of the software and learning procedure.

Previous work on usability evaluation can be categorized as follows:

1. Empirical testing proposed by Molich and Nielsen, (1990)
2. Inspection proposed by Nielsen and Philips, (1993)
3. Comparative Usability measures presented by Dillon and Maquire, (1993)
4. Formal complexity measures introduced by Thimbleby, (1994)
5. The MUSIC methodology is introduced by Bavan and Macleod, (1994)

Each of these approaches has advantages and some have serious weaknesses as well-explained by Chang et al. (1996). Usability evaluation has two particular objectives:

i. Isolate specific problems with the user interface
ii. Obtain an overall figure of merit of the usability for a user interface

Most of the previous approaches concentrate on the first objective. In this thesis, we will concentrate largely on the problem of obtaining the overall figure of merit of the user interface both for all aspects of the user interface together and on figures of merit for different aspects of the user interface.

As it is discussed by Squires et al. (1999), they stated that the affect of usability features on educational context to accomplish educational targets is not considered completely by investigators. For this purpose, authors come up with the necessity to assist the researchers in usability and interactive learning perspectives.

The most interesting challenges take place between designers and HCI investigators nowadays for developing software has the ability to combine both new users and supporting their learning process although having long distance issues. Beside User-Centered Design (UCD) providing high usability development methods and tool accessibility required by ISO 9241, (1998), they stated the demand of the Learner-Centered Design (LCD) is discussed by Quintana et al. (2001) having the ability to create novel learning domains in terms of education purposes would be more effective. LCD has the advantage of placing the user at the center of the educational software design as an online user, putting forward the significance of the learner in the system.

From the e-learning purpose, we should definitely distinguish and clarify that the expected user performance using the software is not just related with completing tasks but it should be strongly connected to learning while completing the tasks, so the key issue is learning. In the design process of the learner-centered applications we should focus on competence and proficiency level of the users in terms of their knowledge, talents and comprehension. Furthermore, separating the users and the learners is an important phase due to their existing differences realization. During the transition from UCD to LCD, learner needs should become definite and differed from the typical system users.

The importance of the instructional interface arises in course of the system user who is referred as learner, intends to learn the concepts instead of aiming to access concepts. So the expressed idea in this context is, providing users to rivet their objectives not only the tasks. It is obvious in the literature that the usability is a very potent quality criterion for the website design to identify user satisfaction by using the application. Because of this, it can be easily stated that usability is a very significant measure affecting the learning.

2.2.1 Usability and E-Learning

International Organization for Standardization (ISO), (1998) defined usability as the measurement of a product that can be used by particular users in order to achieve specific objectives with effectiveness, efficiency and satisfaction in a particular context. In Web based learning systems student satisfaction of using software may be of crucial significance in attaining pedagogical goals.

Finding out learner preferences and adjusting educational environment to his/her needs are the best ways to ensure the learner approval. However, tailoring software into every student requirements may be costly and difficult to achieve. Building learner groups of similar usability preferences can be a good solution to the problem.

Usability of Web based educational systems has become an object of research in recent years. Kukulska-Hulme et al. (2004) stated that good usability depends on such factors as user satisfaction and enjoyability. Squires et al. (1996) emphasized the importance of fulfilling learner requirements.

Many authors considered software evaluation as a very important source of information for identifying student needs. Quinn et al. (1999) presented a technique of evaluating e-learning systems, which takes into account design factors together with acceptance factors. The first ones were investigated by Marsico et al. (2004) described website design issues with all the features that influence website usability.

Usability needs of different groups of users are often connected with color preferences. An appropriate application of colors can make interfaces easier to understand and use. Currently, colors do not often address user needs and only few authors consider user oriented color models. For example, the research of Noiwan et al. (2006) should be mentioned: they discussed color and visual attention of animated graphics for two cultural groups: American and Thai.

Beaudoin, (2006) stated that a necessary element for online course efficiency is establishing a collaborative learning environment depending on student profiling. It was shown that learner usability needs differ depending on learning style preferences of learners. A study of building a customized interface in an intelligent learning environment related to learners' preferences connected with their learning styles was presented by Cha et al. (2006).

From the user's perspective who intended to use interactive software, usability is the most significant property of the software. ISO 9241, (1998) explains the usability as the measurement of a product that can be used by particular users in order to achieve specific objectives with effectiveness, efficiency and satisfaction in a particular context.

Usability strongly affects the success of the e-learning applications and plays a critical role. Let's think if the e-learning application is not easy to use, this will cause users to spend so much time to discover to solve the system functionality instead of working on the learning process which is the main focus of the developed software. Furthermore, if the developed software some problems related to its design issues such as working slowly, inflexible and inelegant it is most probable that users will not eager to use it.

This is a very compelling way to make the users spend very long time on discovering how the system works based on poor designed software in place of understanding the learning context will be distracting and disturbing for the users to figure out the real focus point of learning.

According to Norman, (1993) determinative product should include the following components:

i. Easy to interact with and obtains feedback
ii. Has special objectives
iii. Motivate, interacting with the system to complete the learning objectives
iv. Provide convenient tools
v. Deflect distractions and elements breaking in the learning process

In addition, these components constitute a meaningful structure for pedagogical aspect. By indicating that the learning tools help with users learning process by defining user specific tasks development, rather than just using and acting to use the new technology.

Also these components should be very well-designed and suggests different learning processes and methods from the pedagogical aspect.
For instance, the usage of multimedia must be attentively designed in order to prevent an adverse effect on overload of audio channels. It does not strictly mean that choosing to use the brand new technology is rejecting traditional techniques like simulated systems and problem solving. Instead we can combine all the new and traditional techniques and use this new integrated format in order to perform better teaching.

By considering the feature set of the e-learning applications, we can infer that they have strong effect on organizing the content and providing better functionality and navigation, simplicity of the usage and having more personalized data.

The most important point to develop a system is using the usability criteria while adopting the methodology of Learner-Centered Design Approach. In the User-Centered Design (UCD) approach we mostly deal with users belonging to the same culture and having similar experiences from the application domain perspective, on the other hand in Learner-Centered Design (LCD) different users' clustering must be presented under the consideration of learning strategies, different experiences and motivations to complete the assigned task.

In most cases, users of the system do not have much experience in learning field. Furthermore, they have either no idea about in which learning field they are operating or have very small knowledge or yet having wrong perception. While trying to achieve an assigned task, users mostly do not have the same motivation; what's more their motivation will be effected by success rate by depending on their experience in the related learning domain. At this point, we should introduce to provide the continuous help to the users and give much importance to the necessity of helping and approving of their improvement.

We can describe the main objective of the LCD as filling the gap of observation between the users and learning field and make sure that the users obtain all the information from the related learning material. Furthermore, from the LCD side we are doubling the learning problem with forcing the user to practice a tool which he or she does not use before and learn a material that he or she not been taught before.

As we remarked previously, without taking into account technological subjects, we should be aware of relying upon the educational theory to persuade the developers to construct convenient applications. Contemporarily, constructivist theory is taken up seriously worldwide. This theory states that learning is mostly related with personal genesis, shaped by getting affected by various different essences to reach better knowledge and understanding and refinement of the learning material.

This causes that the assistance should be provided to the users while they are constructing and refining the learning subjects by interpreting in an essential ways. From the constructivists' aspect, we should encourage the users for being charged of their learning and to be experienced by their successfully achieved learning process.

As an extension to the constructivist approach by social aspects in the learning period, a new terminology called situated learning is introduced. This new kind of learning denotes that learning is highly influenced by how communication technology and information is used under the learning environment. By meaning of the learning environment we intended users and artifacts collaborating with the system and making contribution to learning duration. By blending the constructivism and situated learning a new concept called "socio-constructivism" is comprised by Squires, (1999).

With the help of all these considerations, it turned out that not only usability but also effective learning issues should be taken into account for evaluating the e-learning application. For this study as a long term objective, we want to come up with a good definition of the evaluation methodology which addresses both issues. The main focus of the work represented in this thesis is the usability aspect of e-learning software development.

There are different existing evaluation techniques having tradeoffs such as effectiveness and cost while trying to select through them. For some of the methods such as heuristic evaluation, it is easy to control and having low cost but on the other hand if you intend to use them you will face with many problems. These caused problems are mostly arising in the application phase, because we intend to apply small set of rules to a wide

range system. This is figured out by miscellaneous investigators developing specific guidelines for particular applications.
As Notess, (2001) pointed out that for the usability testing special consideration is needed with the help of web-based learning platforms. There have been several proposed ideas tried to meet this requirement. Dringus, (1995) proposes that the usability heuristics of Schneiderman, (1997) and Nielsen (1993) are highly applicable for evaluating e-learning software development. Ravden et al. (1989) stated that the required functionalities of the system should meet are visual clarity, consistency, flexibility and control. According to Schwier et al. (1993) the required principles listed as consistency, time, simplicity, clarity, usage of the white space and minimum memory load.

Squires et al. (1999) adopted the Nielsen's heuristics, (1994) in their work by stating the socio-constructivism as matching of the designer and the user. They proposed to support the learning process in a personalized way of the user by using the level of the user control and matching with the curriculum. But these studies are not really fortified with conducted studies.

Our approach can be identified to evaluate the usability by using a methodology resolving some of the drawbacks of the heuristics and design the study of observers. This methodology has been adapted to the e-learning context.

Course content, as the major part of every E-learning system plays a significant role in success of E-learning organizations and affects the quality of learning. Therefore in addition to considering factors such as user satisfaction, memorability, learnability, error prevention, and efficiency which affect the usability of a system from the interface view point, the factors which associate with usability of the course content should also be considered in designing and evaluating E-learning systems.

Technical issues also should be considered:
E-learning content must be a meaningful to engage learners. Hence technical issues should be considered in order to design effective content. Technical issues such as customization, low bandwidth requirement prevention, consistency, and etc. have an important role in engaging learners to the content. Organizations and also educational

institutions which develop e- learning content should also consider reusability, interoperability, durability and adaptability of the e-learning content.
It's also worth mentioning that there should be a balance between the use of technology pedagogy. In fact the key focus must be on the instructional design. Technical issues are not a limiting factor since we have the ability to conceive how they can be used. In other words, technical challenges will not increase the usability of the content unless pedagogical issues and the other dimensions are considered as well.

Instructor capabilities are another issue should be discussed. In e-learning, same as traditional environment, instructors are the key players. The capabilities of instructor inside the environment affect the usability of the course directly. In general, instruction involves the transmission of declarative or conceptual knowledge such as facts, concepts and principles.

2.2.2 Systematic Usability Evaluation

The set of methods that the estimators go through for usability in terms of application are intended to be usability inspection. Furthermore, these estimators obtain predictions by using their experience. If the other evaluation methods are considered for usability, for instance user-based evaluation, the usability inspection methods are desirable due to their cost-effectiveness. Also, they do not have the necessity for laboratory facilities to save the interactions of user or costly domain analysis, or pervasive unprocessable interviews.

It is asserted that the method for usability inspection "save users" is proposed by Jeffriess et al. (1991) because of not requiring special tools and laboratory equipments. At the same time, these methods contain the most worthy and reliable resource of usability challenges statements. Yet, methods are vigorously connected with the inspector abilities and experience; hence it may appear that distinct inspectors generate distinct results. Identifying a prevalent framework for usability evaluation is the main objective of the SUE (Systematic Usability Evaluation) is proposed by Matera et al. (2002). SUE stated that in order to reach a reliable evaluation, the idea should be kept track of has to be the systematic combination of both inspection and user-based evaluation.

It is specified in couple of studies that how these techniques are completing each other and how they can be used in order to provide a reliable evaluation period. By satisfying the consistency in these studies, matching inspection activities and user tests are offered to be performed by SUE and accurately described how to integrate them to achieve more certain and cost effective evaluation.

The pivot functionality of the inspection is guaranteed by the presence of the experienced evaluators examining the application. In addition, the reason to apply user testing may be coming across with critical situations. For example, it might be required by the evaluator that assuring more intensive evaluation that provided with the help of included users. In this manner, users can concentrate on more decent and resources of the user are structured in an organized fashion. Finally, the total evaluation becomes less expensive but yet powerful.

Many of the current approaches concentrate on particularly demonstration aspects of the utilized interfaces in the application which are pervasive among all interactive implementations related to evaluating the system usability. For instance, the lay-out design, icon selection and interactivity manner, error handling technique consider usability evaluation.

As SUE suggested that the analysis of the implementation should be conducted from different perspectives with the help of particular measurements. Interactivity and demonstration characteristic intend to the most general point of view common to all interactive applications. More specific dimensions address the appropriateness of design with respect to the peculiar nature and purpose of the application.

As we said, the SUE methodology requires to firstly identifying a number of analysis dimensions, suited for the class of applications at hand. For each dimension, general usability principles are decomposed into finer-grained criteria. By considering users studies and the experience of the usability experts, a number of specific usability attributes and guidelines are identified and associated to these criteria. Then, a set of ATs (Abstract Tasks) that addresses these guidelines is identified. ATs precisely describe which objects of the application to look for, and which actions the evaluators must perform in order to analyze such objects.

In order to guide the inspector's activity, the inspection is based on evaluation patterns, called Abstract Tasks, addressing the identified guidelines. ATs precisely describe which objects of the application to look for, and which actions the evaluators must perform in order to analyze such objects. They are formulated precisely by means of a pattern template including a number of items.

ATs are formulated precisely by means of a pattern template including a number of items:

i. AT Classification Code and Title univocally identify the AT, and its purpose.
ii. Focus of Action lists the application objects to be evaluated.
iii. Intent clarifies the specific goal of the AT.
iv. Activity Description describes in detail the activities to be performed during the AT application.
v. Output describes the output of the fragment of the inspection the AT refers to.

During the inspection, evaluators analyze the application by using the defined ATs. In this way, they have a guide for identifying the elements to focus on, analyzing them and producing a report in which the discovered problems are described. According to the SUE methodology, the ATs for evaluating such applications were derived from these usability attributes and guidelines, and described.

2.2.3 Interacting With An E-Learning System

By taking a look at the literature we can state that usability studies in the e-learning field are not very often and very common. Accordingly, the need reveals an all inclusive usability work and the application of a proficient evaluation methodology for the purpose of devising again the user interfaces of the developed applications. Under the consideration of various types of usability evaluation method, we should take account of quantifying the usability while performing the task. These tasks should be predefined and the approach including users performed steps throughout the application interface should be determined.

Due to the fact that usability, as a quality of use in context, is related to the process of use, the steps the user must take and how he achieves the results, the usability evaluation of the systems developed for e-learning is accomplished by testing usability with real users.

These objectives stress the need to quantify usability in terms of user performance and satisfaction. They are measured by the extent, to which the intended goals of use are achieved, the resources that have to be expended in order to achieve them, and the extent to which the user finds the use of the e-learning system acceptable. Such testing is understood to be a combination of behavior and opinion based measures with some amount of experimental control, usually chosen by an expert. It affords information about how users (teachers or students) use the system and identifies the exact problems with a particular interface being evaluated by Zaharias et al. (2005).

To analyze the crucial features of interacting with an e-learning system, we performed a study in which 9 students of undergraduate course at the University of Bahçeşehir have been observed while they were interacting with a DL (Distance Learning) system. The students were observed in the laboratory environment of the Department of Computer Engineering of the University of Bahçeşehir. Student's main objective was to learn some new subjects about the course by using the distance learning system via Internet, and also performing some tasks related with the course such as online tests to appraise the learned content.

Several tools for communication have the capability of exchanging information in between and providing help when needed. Interviews are made for further information collecting from the system users about their experience of using the e-learning software application. Also additional information to improve the existing system gathered and revision of the system can be applied to reorganize the provided services, material and communication tools such as mail, chat and forums.

By making these interviews by using the think-aloud method, it comes out that there are some problems should be solved such:

1. Number of users pointed out that the difficulty of using the software application when the first time they are using a service or progress a new learning path. Furthermore, they complained about the bewilderment and disorientation.
2. Most of the users reported the lack of the highlight mechanisms for important materials and topics covered during the assigned task completion.
3. Mostly, users connected the system from a wrong educational section. This turns out that the need of easy navigation map during the material learning has a strong effect on having better e-learning system usability. Moreover, the system content will be accessible for the personalized users.
4. Users complained the difficulty of searching the relevant material while they perform the learning. This should be simplified by constructing key-word search availability for each material that the user is going to learn.
5. Due to the disconnection problems occurrence during the learning process users obliged to perform all the steps they passed from the beginning. So, users commented that the used software application should have the capability of offline usage.
6. Users stated the self-assessment tools are very motivating in the learning process and provided them to control the improvement in their progress.
7. Users indicated the help they got by using the communication tools which establish collaboration in learning. Whether synchronous and asynchronous system interactions can be performed by multiple users in their learning process.

By this study we can approve the complexity of the e-learning system usability. Importance of presenting the material should be taken into account from the indication of help aspects. Furthermore, the existence of the hypermedia tools provides the capability of personalizing the learning period and connecting through distinct channels, effective adaptation.

Finally, user’s initiative should be encouraged: the participants preferred self-assessment tests to evaluate their progress. The above aspects are related not only to the e-learning environment, but also to the structure of the educational material.

2.3 HEURISTIC EVALUATION IN E-LEARNING

Heuristic evaluation (HE) is a usability inspection technique originated by Nielsen, (1992) and (1994 in which limited number of experts evaluated, followed the set of usability heuristics (principles), to determine if an application is appropriate for those principles and detect the usability problems in the development. The one used most widely is usability evaluation method UEM for computer system interfaces. It is described as fast, inexpensive, and easy to perform, and can result in major improvements to user interfaces proposed by Belkhiter et al. (2003) and Blandford et al. (2004) .

HE is often done during development, but can be very effective when used on real, operational systems as in this study, so as to compare the results of evaluation by two different methods on exactly the same application.

Over the past decade there has been widespread use of web-based e-learning applications for distance and classroom learning, yet little has been done to critically examine their usability. In the consideration of usability features, they are not mostly think of in the development due to instructors and material developers are not trained to complete them or due to their lack of knowledge in those skills. Another reasons for them is evaluation is difficult, time consuming and also expensive.

Furthermore, heuristic evaluation has low cost attributes and simple and it is very beneficial due to its effectiveness and efficiency and appropriateness to identify the problems that delay the users, it is also very important and fitting method in evaluation of e-learning applications.

2.3.1 Evaluation Criteria and Heuristics

Evaluation criteria/heuristics (terms used interchangeably) for e-learning should address interfaces, usability and interaction from HCI, as well as pedagogy and learning from education. A study was undertaken by the authors to establish an appropriate set of 20 criteria within three categories for evaluating WBL applications. This multi-faceted framework supports Hosie et al. (2001) concept of ‘context-bound’/‘context-related’ evaluation.

Nielsen’s heuristics, (1994) form the basis of Category 1, extended to customize them for educational purposes by using Squires et al. (1999) proposed work “learning with software” heuristics. Each criterion has a list of associated sub-criteria or guidelines to help evaluators. These are also shown, but they can be customized to other contexts. Regarding personal learning experiences, which are specifically for learners/students (terms used interchangeably) but not for expert evaluators, so that 15 of the 20 criteria were common to both studies. Evaluations should be performed in a professional and ethical manner.

Participants’ rights must be protected. End user or expert evaluators should sign a consent form and receive information documents, explaining the purpose, and assuring anonymity and those findings will be used for stated purposes only. These evaluations were conducted in an ethical and socially responsible manner. Participation was voluntary and participants could withdraw at any time. Students received no reward for participating. Expert evaluators were not paid for their professional expertise, but received a token gift in gratitude. Furthermore, where questionnaires are used, they should be pilot tested to avoid any potential misunderstanding.

2.3.2 Usability Attributes in E-Learning System Using Hybrid Heuristics

Heuristic evaluation is a usability engineering method for finding the usability problems in a user interface design in e-learning proposed by Nielsen et al. (1990) and Nielsen et al. (1994). Heuristic evaluation involves having a small set of evaluators examine the interface and judge its compliance with recognized usability principles known as heuristics.

Usability Attribute in E-learning:

Usability is a guideline to be applied when developing a system. It makes the application is learnable and allows user to accomplish goal efficiently and effectively. Another purpose of usability is to ensure user’s satisfaction when using an application.

Another purpose of usability is to ensure user's satisfaction when using an application. An application guided with usability will increase the degree of its usefulness which means it is highly usable as the contents are relevant and meaningful to the user.

While a large number of organizations have adopted e-learning programs, far fewer have addressed the usability of their learning applications. It is important to implement usability in E-learning because of several factors. Firstly, is to avoid failure of functionality of the system, if the users are unable to use the system to accomplish their goals. For example, it is considered failure to an E-learning if the user did not know how to download a course's notes.

Secondly, the implementation of usability attributes in E-learning is to ensure it is usable even though it uses simple design rather than focusing on the technology side but the users are unable to use the technology effectively. It is also important to implement usability into an E-learning system as it will increase the quality of courses, states clear objectives, increase relevancy of content, and enhance comfort level with technology. The availability of technical support must be included too, as well as the ability to interact with peer learners because learning process uses communication most of the time.

Heuristics:

There are three common types of heuristic applied in many systems. Nielsen's 10 Heuristics of Usability, (1994) is widely used for web based system, Schneiderman's Eight Golden Rules, (1997) is best known for interfaces and Hybrid Heuristics, the most suitable one to be applied in E-learning system.

1) The Jacob Nielsen's 10 Heuristic Of Usability

Jacob Nielsen is the founder of 10 Heuristics of Usability. His guideline is suitable for any web application or websites. The attributes of Jacob Nielsen 10 Heuristics of Usability are visibility of system status, match between system and the real world, user control and freedom, consistency and standards, error and prevention, recognition rather than recall, flexibility and efficiency of use, aesthetic and minimalist design, help user recognize, diagnose, and recover from errors and the last one help and documentation.

2) The Eight Golden Rules of Schneiderman is focusing on interface of a system. The following are the attributes of the eight golden rules:

i. Strive for consistency.
ii. Enable frequent user to use shortcuts
iii. Offer information feedback
iv. Design dialogs to yield closure.
v. Offer error prevention and simple error handling.
vi. Permit easy reversal of actions.
vii. Support internal locus of control.
viii. Reduce short-term memory load.

To improve the usability of an application it is important to have a well designed interface. Shneiderman's "Eight Golden Rules of Interface Design" are a guide to good interaction design. A good interface will enhance the experience, usability, effectiveness or even success of e-learning system.

Hybrid Heuristic:

Another type of heuristics is the Hybrid Heuristics. It is a combination of Jacob Nielsen's 10 Heuristics for Usability with usability attributes that match with E-learning system proposed by Benson et al. (2002). The attributes of hybrid heuristics are as follows:

a. Visibility of system status
b. Match between system and real world
c. Error recovery and exiting
d. Consistency and standard
e. Error prevention
f. Navigation support
g. Aesthetics
h. Help and documentation
i. Interactivity
j. Message design
k. Learning design

l. Media integration

m. Instructional assessment

n. Resources

o. Feedback

2.3.3 Evaluation Criteria for Web-Based Learning

Category 1: General interface usability criteria (based on Nielsen's heuristics, modified for e-learning context):

Visibility of system status:

i. The website keeps the user informed about what is going on through constructive, appropriate and timely feedback.
ii. The system responds to user-initiated actions. There are no surprise actions by the site or tedious data entry sequences.
iii. Match between the system and the real world i.e. match between designer model and user model:
iv. Language usage such as terms, phrases, symbols, and concepts, is similar to that of users in their day-to-day environment.
v. Metaphor usage corresponds to real-world objects/concepts, e.g. understandable and meaningful symbolic representations are used to ensure that the symbols, icons and names used are intuitive within the context of the performed task.
vi. Information is arranged in a natural and logical order.

Learner control and freedom:

i. Users control the system.
ii. Users can exit the system at any time even when they have made mistakes.
iii. There are facilities for Undo and Redo.

Consistency and adherence to standards:

i. The same concepts, words, symbols, situations, or actions refer to the same thing.
ii. Common platform standards are followed.

Error prevention, in particular, prevention of peripheral usability-related errors :

i. The system is designed such that the users cannot easily make serious errors.
ii. When a user makes an error, the application gives an appropriate error message.

Recognition rather than recall:

i. Objects to be manipulated, options for selection, and actions to be taken are visible.
ii. The user does not need to recall information from one part of a dialogue to another.
iii. Instructions on how to use the system are visible or easily retrievable whenever appropriate.
iv. Displays are simple and multiple page displays are minimized.

Flexibility and efficiency of use:

i. The site caters for different levels of users, from novice to experts.
ii. Shortcuts or accelerators, unseen by novice users, are provided to speed up interaction and task completion by frequent users.
iii. The system is flexible to enable users to adjust settings to suit themselves, i.e. to customize the system.

Aesthetics and minimalism in design:

i. Site dialogues do not contain irrelevant or rarely needed information, which could distract users as they perform tasks.

Recognition, diagnosis, and recovery from errors:

i. Error messages are expressed in plain language.
ii. Error messages define problems precisely and give quick, simple, constructive, specific instructions for recovery.
iii. If a typed command results in an error, users need not retype the entire command, but repair only the faulty part.

Help and documentation:

i. The site has a help facility and other documentation to support users' needs.
ii. Information in these facilities is easy to search, task-focused, and lists concrete steps to accomplish a task.

Category 2: Website-specific criteria for educational websites:

Simplicity of site navigation, organization and structure:

i. The site has a simple navigational structure.
ii. Users should know where they are and have options of where to go next, e.g. via a site map or breadcrumbs.
iii. The navigational options are limited so as not to overwhelm the user.
iv. Related information is placed together. Information is organized hierarchically, moving from the general to the specific.
v. Common browser standards are followed.
vi. Each page has all the required navigation buttons or hyperlinks (links), such as previous (back) next and home.

Relevance of site content to the learner and the learning process:

i. Content is engaging, relevant, appropriate and clear to learners using the WBL site.
ii. The material has no biases such as racial and gender biases, which may be deemed offensive.
iii. It is clear which materials are copyrighted and which are not.
iv. The authors of the content are of reputable authority.

Clarity of goals, objectives and outcomes:

i. There are clear goals, objectives and outcomes for learning encounters.
ii. The reason for inclusion of each page or document on the site is clear.

Effectiveness of collaborative learning (where such is available):

i. Facilities and activities are available that encourage learner-learner and learner-teacher interactions.
ii. Facilities for both asynchronous and synchronous communication, such as e-mail, discussion forums and chat rooms.
iii. Level of learner control:

iv. Apart from controlling the interactions with the site, learners have some freedom to direct their learning, either individually or through collaborative experiences, and to have a sense of ownership of their learning.
v. Learners are given some control of the content they learn, how it is learned, and the sequence of units.
vi. Individual learners can customize the site to suit their personal learning strategies.
vii. Educators can customize learning artifacts to the individual learner, for example, tests and performance evaluations can be customized to the learner's ability.
viii. Where appropriate, learners take the initiative regarding the methods, time, place, content, and sequence of learning
ix. Support for personally significant approaches to learning:
x. There are multiple representations and varying views of learning artifacts and tasks.
xi. The site supports different strategies for learning and indicates clearly which styles it supports.
xii. The site is used in combination with other mediums of instruction to support learning.
xiii. Metacognition (the ability of a learner to plan, monitor and evaluate his/her own cognitive skills) is encouraged.
xiv. Learning activities are scaffolded by learner support and by optional additional information.

Cognitive error recognition, diagnosis and recovery:

i. Cognitive conflict, bridging and problem-based learning strategies are used in the recognition-diagnosis-recovery cycle.
ii. Learners have access to a rich and complex environment where they can explore different solutions to problems.
iii. Learners are permitted to learn by their mistakes and are provided with help to recover from cognitive errors.
iv. Learners are given opportunities to develop personal problem-solving strategies.

Feedback, guidance and assessment:

i. Apart from the system's interface-feedback by the system, considered under Criterion 1, learners give and receive prompt and frequent feedback about their activities and the knowledge being constructed.
ii. Learners are guided as they perform tasks.
iii. Quantitative feedback, e.g grading of learners' activities, is given, so that learners are aware of their level of performance.

Context meaningful to domain and learner:

i. Knowledge is presented within a meaningful and authentic context that supports effective learning.
ii. Authentic, contextualized tasks are undertaken rather than abstract instruction.
iii. The application enables context- and content-dependent knowledge construction.
iv. Learning occurs in a context of use so that knowledge and skills learned will be transferable to similar contexts.
v. The representations are understandable and meaningful, ensuring that symbols, icons and names used are intuitive within the context of the learning task.

Learner motivation, creativity and active learning:

i. The site has content and interactive features that attract motivate and retain learners, and that promote creativity on the part of learners, e.g. the online activities are situated in real-world practice and interest and engage the learners.
ii. To promote active learning and critical thinking, tasks require learners to compare, analyze and classify information, and to make deductions.

3. LITERATURE REVIEW

The consideration of the usability from the user's context is dealing with applications which provide simple learning, efficient utilization and entertaining. The ISO 9241, (1998) standard has defined usability as the measurement of a product that can be used by particular users in order to achieve specific objectives with effectiveness, efficiency and satisfaction in a particular context

With this indicated standard, application's usability can be reached by three main feature measures which are: effectiveness (the criteria indicated the ability level of the user interacting with the system in order to accomplish the task he or she is assigned. Efficiency is another measure related with the task completion speed of the user and finally the third and the last one is satisfaction which represents the level of the pleasure that the user has when interacting with the application.

3.1 USABILITY IN THE E-LEARNING CONTEXT

Developed supportive learning computer applications belonging to cognitive domains should be take into consideration as very distinct applications when compared with mercantile operation functioning implementations. The latest version of the applications has the capability to support rapid task completion and prevent extended and repeated fulfillment operations. Adversely, e-learning applications' objective is to support learning process by using knowledge transmission not knowledge translation which helps to direct interaction in the education and obtains additional support for intellectual thinking during the learning process. Besides knowledge transfer helps implementing the attitude alteration and decreases the interest to the technology with requiring high cost of the learning the material.

There is an absolute difference between these two types of implementations which may require manipulation to the conventional interaction for referring exclusive usability requirements for e-learning aspect. For example, there should be a relaxing way for users while they perform learning by discarding some of parts in the material which are not related to their learning. The form of mercantile operation-based implementations apprehend for user-centered design. Furthermore, designers can be inclined to think the uniformity aspects between system users whereas e-learning concentrates on learner-centered design.

The objective of e-learning applications is identifying nonhomogeneous learning user sets consisting of different learning styles, encouragement, background for computation and level of experience. Furthermore, application should satisfy the distinct usability needs of the system users. The interface of the implementation should be well configured; in this manner it obtains trivial and effective navigational techniques. Additionally, application interface provides customized learning material matching with the users' assigned tasks. Also, e-learning systems are far beyond of being educational software but also they should be considered as computer systems as well.

The users of the system should be experienced before starting to perform assigned tasks related to the effective learning. But in this phase, it is not desired that users going to spend large amount of time trying to figure out how to use the application. The interface of the application should be perceptive, by this way even a new user joins the system he or she can interact with the system promptly.

By leaving out the classical usability measure, we should take into account that e-learning systems should be appraised for pedagogical capability. This capability for usability supports various types of the users with different learning methods by depending on the learning goals. Furthermore, pedagogical capability provides suitable tools for users such as interfaces, materials and user tasks.

In the evaluation part, it is not reasonable to devote usability and pedagogical capabilities due to their relevance. Stating the matter differently, usability cannot be thought of separately like not considering the material and functionality of learning in the evaluation of the e-learning application. But we can indicate that there are extreme differences between UT of classical assignment-based implementations and UT of distance learning.

It is not suitable for e-learning, if we use low accomplishment time of tasks as an argument leads to implementation efficiency because for each user of the system there are various directions for learning. Furthermore, low accomplishment time is not a very good measure for usability and the approach for dealing with errors is totally different.

There is a significant difference between usability errors and cognitive errors in terms of e-learning. Quality of learning can be ruined by usability errors so they will be avoided whereas cognitive errors are allowed because they constitute learning process components and include higher order consideration.

There is a general belief that people learn from mistakes and the best learning depends on correcting those mistakes. Nevertheless, the external help for the learner in the moment they faced with a problem during the learning should be provided and save the user from potential errors.

3.2 EVALUATING E-LEARNING USABILITY

It is widely approved by many researchers [58, 1, 26, 11] that e-learning should have the uniqueness property. Various studies proposed to evaluate this. For example, Squires et al. (1999) evolved a set of standards called learning with software heuristic, related with Nielsen's, (1994) approach in heuristics and socio-constructivist learning technique. By using these heuristics, teachers will have an effective tool for improving e-learning application quality in terms of visual usability and effectiveness issues.

In Masemola et al. (2006) proposed work; they examined the UT of e-learning implementation for UT tasks, steps and procedures aspects. They conducted their tests in a laboratory environment in the proposed study, by assigning tasks to the users to follow the e-learning manual named as Relations.

It is found out that time usage pattern metric is a useful metric helping to distinguish time spent navigation and time spent on learning materials. Furthermore, data analysis leads to figuring out the different learning characteristics. Application and refinement on single user and co-discovery tests can be considered as an appendage to the existing Masemola and de Villiers's UT framework, (2006) for e-learning.

This study has provided numerous principles and guidelines that can steer designers in taking their decisions. Nevertheless, it is no substitute for system assessment. To enable and facilitate design according to usability engineering principles, usability evaluation plays a fundamental role in a human-centered design process. Usability as a quality of use in a context should be viewed as comprising two essential aspects:

1. Efficacy in use, considered primarily as involving measures of user performance, and
2. Ease of use, considered primarily as involving subjective judgments.

Although usability is the basic parameter for the evaluation of e-learning technologies and systems, the idea of e-learning usability is still quite new. For this reason, the design of accessible and easy to use e-learning system able to address the needs of all potential users requires additional considerations.

Comprehensive research concerning an evaluation methodology is needed. The assessment presented in this thesis concentrates on the system usability aspects. The usability evaluation techniques employed in the assessment of onsite and web-based systems, as well as the evaluation procedure and the results achieved, are described.

4. METHODOLOGY

This research has two folds in the methodology part:

a. Research of action
b. Procedural Model as Masemola and de Villiers specified in their framework

Research of action mainly takes place on evaluation purposes and applied involvement of the investigators. Mostly, several investigators from different stages are getting involved at this research of action process.

Procedural model is mentioned in Masemola et al's framework, (2006) basically targets the user tasks under the interactive e-learning systems being a part of cognitive domains.

If we compare these two design types while considering our study, we can state that action research is cyclic because of its nature, and involves multiple iterations. Investigators are required to actively participate and collaborate with other stakeholders. While performing data collection and analysis, researchers will be the key people to make a decision.

Furthermore, for the generated data by the researchers we can specify that this data contains more quality rather than quantized data. After conduction of the analysis based on planning and monitoring, these are used as feedbacks for future iterations and process is going to be completed while the result is reached the bottom line.

Masemola et al's framework (2006) comprises these steps as follows:

i. Objectives set up
ii. Decision on required measurements and metrics
iii. Formulation of the necessary documents

Third step contains various types of document collections which are: task list, participant's information document, test conductor's checklist, test plan and post questionnaire in order to measure user's satisfaction after using the e-learning system.

iv. User empowering
v. Pilot test application
vi. Adjustments on documents such as task list, user information and test plan based on using the feedback of pilot test application
vii. Usability testing
viii. Identification of the analysis and stating the uniqueness of conducted test by e-learning aspects.
ix. Obtain results and recommendations

4.1 USABILITY METRICS

In order to measure the user performance and quality, quantitative data analysis conducted while forming the usability metrics and to measure the perception of the users dealing with the system subjective data analysis conducted.

In this study, performance measures are considered as follows:

a. Number of user errors
b. Number of errors repeated by user
c. Number of mouse clicks
d. Number of user's call for help
e. Time to complete assigned task to user
f. Time to learn to the given course material
g. Time to read the material
h. Number of answers given correct by the user

After forming these quantized metrics, we have conducted a post-questionnaire to each users of the system for further analysis.

4.2 PROBLEM DEFINITION

Evaluation of the usability of e-learning could be used to isolate learning problems with particular e-learning packages and propose methods for correcting these and/or to obtain an overall figure of merit of the usability of the system to allow comparisons between different-learning systems. The focus of the usability measurement approach discussed in this paper is on obtaining the figure of merit for all aspects of the e-learning system and for each different aspect of the e-learning system.

Usability of an e-learning system is a difficult concept to quantify as it involves fuzzy concepts and multiple dimensions and factors that could be subjective. When examining usability, Dillon et al. (1993), isolated four major aspects (a) user performance (b) user attitude to the system (c) mental effort required and (d) formal analysis of the system to obtain operational complexity. Each one of these dimensions themselves can be further decomposed. User performance can be measured by time taken to complete the task; it could also be measured by user efficiency or relative efficiency which is the ratio of user efficiency divided by expert efficiency.

Similarly, measurement of user's preference and attitude can itself be decomposed as described in the Metrics for Usability Standards in Computing (MUSIC) system proposed by Dillon et al. (1993) and Nielsen, (2000). So when one examines each of these, one notices that there are several factors which go into the mix for characterizing usability.

In this work, the basic problem we intend to address is to isolate the most important of these factors and work out a means of characterizing each of them individually. And then to develop a technique which seeks to combine them in such way that we obtain a composite overall figure of merit for the total e-learning system, as well as a figure of merit of the first three dimensions referred to above, namely the actual user performance, user preference and cognitive load.

The difficulty in combining these factors lies in the fact that each one of these factors on their own is an imprecisely defined concept and has fuzzy aspects. Furthermore the notion of usability itself cannot be readily quantified into a numerical value, which can be immediately understood as a figure of merit. The notion of usability itself has a number of fuzzy aspects. For these reasons we have chosen to examine fuzzy systems theory as a mechanism for characterizing each of these individual factors and dimensions and also for characterizing usability itself. In the rest of this paper, we will design a measure of usability and measures of each of the dimensions of the usability based on fuzzy system theory. This fuzzy system based usability measure can be utilized in the prototyping and implementation stages to characterize and improve the user interface.

4.3 FUZZY SYSTEM BASED USABILITY METRICS FOR E-LEARNING

There are several studies consider the usability of a system by using fuzzy approach. At first we will classify these studies according to their applications and applied methods. Table 4.1 indicates the existing fuzzy logic used systems.

Table 4.1: Fuzzy Application Classification

Researcher	Date	Application	Method
Dubey et al.	2012	Measurement of Object Oriented Software Usability using Fuzzy approach	ISO 9241 as a base model, metrics used: effectiveness, efficiency, satisfaction and newly learnability added
Yahaya et al.	2011	Software usability using Fuzzy Approach	Fuzzy set conjoint model development based on usability and satisfaction
Lu et al.	2009	Usability Measurement Using a Fuzzy Simulation Approach	FSM simplification by reducing the redundant states, and propose a fuzzy mathematical formula for calculating the probability of state-transition
Kumar et al.	2010	A Fuzzy Logic Approach to Measure Complexity of Generic Aspect-Oriented Systems	In order to automate complexity measurement, a tool has been developed using fuzzy logic, in which some set of rules have been defined as rule base. Using this tool, complexity of majority of AOP languages can be measured,
Chang et al.	1997	An Intelligent System Based Usability Evaluation Metric	Development of a model for each of the dimensions using fuzzy set theory. Supported with Takagi Sugeno fuzzy inference approach developing the overall measuring of usability.

To describe the usability of an e-learning system implemented metrics are as follows:

a. Web System Feedback
b. Error Prevention
c. User Performance
d. Consistency
e. Error recovery
f. Privacy
g. User like/ dislike
h. Internationalization
i. Cognitive Load
j. Help

In order to talk about the measurement of these usability metrics of the e-learning application listed above we will discuss each of them in the following section.

1. System Feedback:

The system should always provide users with the appropriate feedback so that they can be informed about what is going on in the system at any time.

System feedback is characterized by a number of aspects including:

i. Error localization
ii. If an action is not permitted, does the system say why?
iii. Does the system give prompts as to what to do next?
iv. Does the system let one know where one is?
v. Does the system tell one why one cannot do something?

Inadequate system feedback has several components and these include:

a. No. of times dialogue/feedback is missing
b. No. of times dialogue/feedback unnecessary
c. No. of times system feedback confuses the user (information, tasks, path, message, help)
d. No. of messages that are irrelevant

e. No. of actions taken which lead to repeat of the same system feedback message
f. No. of times the user makes the wrong choice of action based on system feedback
g. No. of times user grumbles about the feedback messages

Each time anyone of these components is not satisfied, we can say that on this occasion system feedback was inadequate. Thus we will count this as contributing one unit to a measure called inadequate system feedback. System feedback itself is usually characterized using linguistic terms such as Good, Average, Poor. The Fuzzy Term Set therefore is {Good, Average, Poor}. If we employ triangular shapes for the membership functions, the Fuzzy membership feedback will look like Fig. 1a.

It is useful to discuss each of these aspects of inadequate system feedback a-g in more detail in order to understand exactly what they mean and how they are gathered. We therefore discuss them in turn here. During the test, the tester expects some feedback at various times. If the system does not provide feedback when the tester expects it, we take the situation as contributing one unit to the overall inadequate system feedback.

An example of this would be when the user wishes to carry out a print. This takes a finite amount of time for the system to spool the file followed by the print. If the system does not provide the system with any feedback, either with a clock or an hour glass indicating that it is in fact processing something or with a message such as print is currently being spooled, the user could believe that the system has hung up.

Unnecessary system feedback can distract the user and clutter the screen. Therefore a count is done on the number of times a feedback message is given by the system when the user finds that it is unnecessary. Again each occurrence of this contributes one unit to the value of overall inadequate system feedback.

The system feedback could confuse the user either because it is not precise enough or the language it uses is not familiar to the user. Furthermore it could also be obscure to the user because its size and location on the screen is such that it does not easily attract the user's attention. He could therefore be led into carrying out unnecessary actions such as choosing incorrect options because it is unclear what he is to do next.

A message could appear which is not relevant to the particular state that the system is in, or particular set of actions that he needs to carry out. This could distract the user and mislead him into believing that he is in a state other than the one that he is currently in. Each irrelevant message contributes one unit to the overall inadequate system feedback.

If the system provides the same feedback message for different user actions, the user could be misled into thinking that the tasks they are currently doing is the same as the previous task, which resulted in the same feedback message. There could, of course, be differences in the tasks, and the feedback messages concerned should reflect these differences. Each time a different user action gives the same feedback message, it is taken to contribute one unit to the overall inadequate system feedback.

If the user, after reading the system feedback, takes an action which is erroneous or wrong, it implies that the system feedback message is confusing the user or it is unclear as to the course of action he should pursue. Each occurrence of this contributes one unit to the overall inadequate system feedback measure.

This component (g) represents the number of times that the user is annoyed by or dissatisfied with the feedback message. It is measured by asking the user to grumble out aloud on each occasion that he is unhappy about the system feedback message. Each occurrence of this contributes one unit to the overall system feedback.

2. Consistency:

The look, feel and behavior of the interface should be consistent throughout the application and also consistent with other applications in the same domain. Most guidelines, such as those referred to earlier in the paper, seek to bring this about. This consistency should be across a variety of issues such as message display methods, colour use, key definition, data entry methods, etc.

If the user interface is consistent it reduces the amount of uncertainty that the user faces when using the interface. It also is likely to reduce the number of erroneous interpretations or actions that the user makes. Consistency of the interface has a number of components and these include consistency with the respect to the following:

a. Message display methods (prompts, warnings, help)
b. Color use (entry form, menu and submenu, foreground/background)
c. Keys definition
d. Data entry method
e. Menu, dialogue and window display methods
f. Menu hierarchy is consistent with the real world
g. Terminology used is the same as in real life in that domain
h. Menu options have to be consistent with Menu Title.

Each time anyone of these is not satisfied between screen to screen displays, we will count this as contributing one unit to a measure called inconsistency. The Fuzzy Term Set is {Good, Average, Poor}, and the Fuzzy membership for consistency will look like Fig. 1b. It is useful to discuss these components of consistency in a little more detail, in order to assist the person collecting information for this factor of overall usability.

Consistent message display methods require that any message in each category gives a prompt, a warning or help, has the same format, and is different from messages in other categories. The format should be similar in presentation as well as color use. This will ensure that if a particular warning comes up, or if a particular help message comes up, the user is immediately aware of the type of message.

The system user interface has to be consistent in the use of color with the respect to the following: a) work space, b) text, c) menu bar, d) tool bar, e) menu text, f) inactive title bar, f) active/inactive border (for programs where multiple applications can be run simultaneously), g) highlighted text, f) background. Inconsistent color use can be irritating to the user as well as distract him and even mislead him as to the significance of something on the screen. Any special or function keys which are used must be defined consistently.

For example, the "ESC" key might be used to abort a function. If, on another occasion the user uses some other key, say an exclamation mark, to abort a function, then the user could be confused and would have to learn the special keys used in a specific situation, rather than knowing that a particular key is used to perform a similar operation in all situations. As mentioned above, the user interface should be consistent with respect to data entry methods and be consistent with respect to menu dialogue and window display methods.

If, for example, a dialogue box is used for data entry, or a pull down menu with selection from a list of files is used, then that approach should be taken throughout the application. The terminology used should be similar to that which the user is likely to find himself using in the real world. A mistake often made is that terminology used reflects the real world of the system developer rather than the real world of the user.

An example of confusing terminology would be the use of the notion default which to the user means that a variable, if not specified, utilizes the default value. However in the banking environment default to a loan adviser would indicate that the loan had not been met. Lastly menu options should be consistent with the menu title. For example, if one has a pull down menu which says "file", it would be inappropriate to put in "date" as a menu option which would allow one to change the date display.

3. Error Prevention:

Error prevention is an important goal of the design of the user interface. If the user interface specifically helps the user to avoid making errors, it increases his efficiency. It will also reduce the level of frustration the user is likely to experience with the user interface and therefore bring about greater acceptance of the user interface by the user.

There are several aspects which need to be taken into account when measuring error prevention and these include the following:

a. No. of errors encountered during task.
b. No. of wrong key strokes/press causing Error Messages
c. No. of times the same key is pressed without the desired response from system
d. No. of extra key presses that are unnecessary
e. No. of times the same error is encountered
f. No of steps missing compared with real world execution

Each time anyone of these is not satisfied, we count one unit towards error occurrence. The Fuzzy Term Set is {Good, Average, Poor}, and the Fuzzy membership for error prevention will look like Fig. 1c.

Most of the above components which make up the measure of error prevention are self explanatory. It is, however, worth detailing a few of them in order to understand the differences between some of these components.

a. is simply a measure of the number of errors that have occurred during a particular task being conducted by the user when carrying out usability testing. If the number of errors is low, the error prevention of the user interface, at least for that task, is good. If the number of errors is high the error prevention is poor.
b. really relates to the user trying to perform an illegal operation which, instead of carrying out the operation which the user thinks will be carried out, leads to an error message.
c. on the other hand leads the user to believe that a certain key press will produce a certain desired response from the system, i.e. he believes that the key is linked to a certain operation by the user interface, however, that key is not in fact linked to that operation but an alternative key might be. An example of that would be to use the arrow keys to move across to the next desired button in the user interface rather than using the TAB key.
d. relates to occasions when the user believes that a key press is necessary but in fact the key press has no effect.
e. relates to the user in fact generating the same error again and again.
f. relates to the degree of closeness between the steps which the user goes through in the user interface and the carrying out of that similar task in the real world.

If there are steps missing, then it is probably more likely that the user would imagine that they are in place and seek to execute them, leading to generate one of the above error components (a) to (e).

4. Performance/Efficiency:

Performance or efficiency is a quality of the user interfaces which characterizes how effectively or efficiently the user can complete his tasks. Performance and efficiency has a number of components and these are as follows:

a. No. of goal/tasks not achieved
b. Time taken for task completion
c. Unproductive period
d. Percentage of task not completed.

Each time anyone of these is not satisfied, we count one unit towards inefficiency.
The Fuzzy Term Set is {Good, Average, Poor}, and the Fuzzy membership function will look like Fig. 1d.

To elaborate on the components which go into making up the performance or efficiency measure, we note that during usability testing the user is given a number of tasks to complete. The first component measures the number of goals or tasks that the user is unable to complete in the time given for usability testing.

The second component actually measures the time taken to complete the particular task. Tasks seem to vary in duration and complexity and therefore just using the actual time taken may not be a good enough indicator, because it does not contain the notion of the length and complexity of the task. In order to develop the measure of the task completion time, we use the ratio of the time taken by the user divided by the time taken by a known expert user. Therefore:

$$Task\ Completion\ Time = \frac{\textbf{Task completion time of the user}}{\textbf{Task completion time of an expert user}} \quad \textbf{(4.1)}$$

For every ten per cent increase of the task time taken by the user over the task time taken by the expert user, we will count as one unit towards the number of inefficiencies involved in task completion time. If the time taken by the expert user was 100 minutes and the time taken by the user in the usability test was 130 minutes then we would take the contribution towards efficiency caused by task completion time as 3 units.

Unproductive period represents time spent by the user not actually working on the specific task or not working towards the specified goal involved in the task. It consists of time that the user spends referencing help, thinking about what to do next and solving problems not directly associated with the task at hand. Unproductive period is defined by reference as the formula:

$$\text{Unproductive Period} = \text{help time} + \text{search time} + \text{snag time} \tag{4.2}$$

The last component consists of the percentage of tasks the user did not actually complete and it is calculated directly from the first measure of goals not achieved and is given by:

$$\text{Percentage of uncompleted task} = \frac{\textbf{Nonachieved goals}}{\textbf{Number of goals expected to be achieved}} \tag{4.3}$$

5. User Like/Dislike:

Unlike the previous measure which measures the manner in which the user interface facilitates user effectiveness or efficiency, the like/dislike factor measures user preference. This essentially indicates the level of satisfaction which the user feels with the system and the user interface. There are in fact two alternative approaches which can be used to measure this user like/dislike.

Approach 1:

i. Count number of times user gives positive comments when testing the system.
ii. Count number of times user grumbles or gives negative comments.

Approach 2:

Ask user to indicate on a scale of 1 to 10 the likeability of the system.

From approach 1, we add one unit to user likes for

i. We subtract one unit for
ii. In approach 2, we use a scale of 1 to 10 and convert it using the membership function into the fuzzy sets.

Fuzzy Term Set therefore is {Good, Average, Poor}, and the Fuzzy membership function will look like Fig. 1e.

6. Error Recovery:

Error recovery is that quality of the system of the user interface which allows the user to exit from a situation that the user did not intend to be in. Users frequently choose the wrong option or enter the wrong data and they are likely to find themselves in an error state from which they need to recover.

The manner in which the system facilitates this error recovery could reduce the time the user spends recovering from this error state. Error recovery consists of a number of components and these include:

a. No. of times the user has to redo the task
b. No of time the user did not continue
c. No of actions taken that do not solve the problem
d. No of minutes (hours) spent on one error recovery
e. Percent of all time spent on error recovery
f. No of times the user has to reboot/start again

Each time anyone of these is not satisfied, we count one unit towards inadequate error recovery. Fuzzy Term Set therefore in {Good, Average, Poor}, and the Fuzzy membership for error recovery looks like Fig. 1f.

The first component essentially means that the user made sufficient errors in a particular task requiring that he has to redo it completely. The second measures the number of times the user has got himself into a condition whereby he is unable to continue the task any further. This state could be the result of the user carrying out an action which was not intended by the system designer but caused the system to enter an error from which there was not an emergency exit to allow the user to continue the task.

The third component measures the number of actions which were unsuccessful that the user takes to get out of an error state. For example if a user finds himself in an error state and he needs to press a certain key to get out of it, if the key is an expected key like ESC he will be able to get out of the error straightforwardly and proceed. If another key was chosen which was obscure and the user was unfamiliar with the user interface, he may end up in the error state trying out a number of keys before he finally hits upon the specific key that allowed him to exit.

The fourth component represents the amount of time the user is actually in the error recovery state. We measure the time in minutes so we count each minute as contributing a single unit to an inadequate recovery. The fifth component measures the proportion of the total time to complete the required tasks that the user spends trying to recover from errors. The last component indicates the system has got into such an error state that the only option for the user is to reboot the system.

7. Cognitive Load:

Cognitive workload is related to the mental effort required by the user to perform tasks using the computer system. If the cognitive workload is high then the user is likely to experience a degree of stress and a degree of pressure in using the particular user interface. If this degree of stress and pressure is unacceptably high, then the user interface needs to be improved in respect to this particular aspect.

There are a series of measures to gauge cognitive workload, some of which are objective and others subjective. The objective measures largely measure heart rate and respiration and they require instrumentation of the user and therefore could be intrusive in actually achieving a proper understanding of the user's cognitive workload, in the sense that the user could react to the electrode attachment rather than just the user interface.

Subjective measures, on the other hand, utilize questionnaires and look at a series of factors to determine the cognitive workload. Two widely used questionnaires are the subjective mental effort questionnaires (SMEQ) which were developed at the University of Kronnigon and Delft University of Technology and the task load index (TLX) which was designed by NASA. The approach explained here could be utilized with either of the above two questionnaires or alternatively with other questionnaires.

We will utilize a count on a series of components to help determine the cognitive workload. However the methodology used here could be easily adapted to be used with any of the subjective measures, or indeed the objectives measures of cognitive workload. The components that we will look at to determine the cognitive workload are as follows:

a. No. of unfamiliar concepts/terminology one has to use
b. No. of unmatched task executions with the real world
c. No. of hours/minutes spent on the first glance of user manual
d. No. of times the manual is used
e. No. of times the user has to access the help window
f. No. of times one has to ask for help from help desk personnel
g. No. of things the user has to remember rather than select on the screen

Each time anyone of these is not satisfied, we count one towards cognitive load.
The Fuzzy Term Set therefore in {High, Average, Low}, and the Fuzzy membership for cognitive load will look like Fig. 1g.

If the user comes across unfamiliar concepts or unfamiliar terminology then he is likely to experience a higher cognitive workload than if he is only working with familiar concepts or familiar technology. The same applies if we come across task execution sequences which do not match with the real world. If the user seems to rely very heavily on the user manual, either in terms of spending a lot of time with the user manual initially before he starts to use the system, or if he has to revert to using the user manual frequently or to access the help window frequently, then the user is experiencing a higher cognitive workload than if he did not have to do these things.

Also, if he has to keep accessing help desk personnel for assistance, then he is in fact experiencing a higher cognitive workload, in the sense that he is unable to directly understand the user interface. If the user interface is such that the user is frequently faced with a selection of items on the user interface rather than having to remember cryptic commands, then the user has a lower cognitive workload. If, on the other hand, he has to remember things like OPTION C for making a copy, or CONTROL C for making a copy rather than just selecting Copy from the menu, then s/he has a higher cognitive workload than in the second case.

8. Internationalization:

One of the differences between an e-learning application and a traditional application is Internationalization. Strength of e-learning is the ability of providing education to remote students who could be from different countries. In order to be able to do this, the e-learning system needs to be designed to handle different languages and cultures. Thus, a good e-learning system should consider the nationality of the users.

Internationalization has a number of components with the respect to the following:

a. Number of times that the user is required to use a language which is unfamiliar.
b. Number of times that the user encounters a situation where s/he doesn't understand what the e-learning system means due to a nationality problem.
c. Number of times that the user's nationality leads to a dislike of the user interface. Each time any one of these is not satisfied, we count one towards low internationalization. The Fuzzy Term Set is: {Good, Average, Poor} and shows the fuzzy membership function for Internationalization.

Point 1 is simply about asking the user to use a particular language to view and/or submit information through the e-learning system. For example, the e-learning system may ask the user to use only English to complete an application form.

Point 2 concerns a situation where the user cannot understand what the e-learning system means because of his/her nationality. For example, if an e-learning system wants to gather some information about sport. The e-learning system may ask "Do you like football?" This question may be confusing to a user from Hong Kong to whom football means "soccer" – a word that has a totally different meaning in Australia. This confusion occurs because of different cultures.

Point 3 is simply about things that the user dislikes about the e-learning system due to his/her culture and/or nationality. For example, people from Hong Kong may not like to see a white and black background as this has funeral connotations.

9. Privacy:

Privacy is one of the important factors on Usability Testing. Low security on privacy or force user to let their privacy to be public will drive users don't like your web system at all. In other hand, if you let users have choice to choose whether they allow web system to publish their information or not, then your site will be much popular then the other. Thus, privacy is important while we build up a web system on Internet.

Privacy has a number of components with the respect to the following:

a. No of time that they require your privacy information as compulsory, which is not really gathering as security or functionality purpose.
b. No of time that the web user interface pops up advertising or non-relative material that is not requires to complete the task.
c. No of time that the web system sends you junk emails, which the web system did not ask for your permission yet.

Each time anyone of these is not satisfied, we count one towards privacy. The Fuzzy Term Set therefore in {Good, Average, Poor}, and the Fuzzy membership for privacy.

a. Although web system used to required user information for security purpose such as surname, date of birth has to be filled before it provide web service. But sometime, web system only attempts to collect marketing information by asking user to fill every single field to proceed further. The manner makes users feel annoying. Thus every time when web system asking some information that are not really relative to security purpose or web functionality need such as password, username etc, we count one toward to low privacy.

b. Pop up advertising frame that is not relative to the task that the user actually performs without asking users' permission will drive user feel annoying. Thus we count one toward to low privacy every time when the unnecessary frame such as advertising.

c. One of the problems that users always complain is not always received junk emails, which they don't even want it from web system. Thus we count one toward to low privacy every time when send junk emails without asking users permission.

10. Online Help:

Online help is one of the essential functions that systems must provide to users so that they can receive assistance when using the system. Thus, a good e-learning system should have a help section that can be accessed by the user as required. Help has a number of components with the respect to the following:

i. Number of times that help is missing within the e-learning page when the user requires help.
ii. Number of times that users cannot find help in the general help sections.

Each time either of these is happens; we count one towards poor help. The Fuzzy Term Set is: {Good, Average, Poor}, shows the fuzzy membership function for Help.

A user would always feel annoyed if s/he could not get help straight away from the on-line learning system that s/he is actually using. Thus, if the e-learning system can provide useful help on what the user needs by s/he clicking on the same page of an e-earning package, s/he will feel happy to use the system.

Conversely, if s/he cannot find help on that page, even if s/he eventually does find help in the general help section, the user may feel that s/he has had to take too much trouble. A user will complain if s/he cannot find help from their chosen page.

4.3.1 Usability Testing

The primary aim of the usability testing for e-Learning system conducted within this project is to generate an extensive set of data, which is to be used to help form the various control rules. The data obtained from the usability testing is to help form the breakpoints of each of the subspaces for each of the usability factors. The results from the test will be used to define the control rules and the breakpoints as accurately as possible, so as to provide a reliable method of determining the usability measure of newly implemented interfaces.

4.3.2 Fuzzy Rule Base For The Usability Problem

As explained above we intend to use the first order Takagi-Sugeno, (1985) approach for fuzzy inference with a linear function for the right hand side. The inputs on the left hand side of the fuzzy rule will consist of the factors or features that affect usability defined in Section above. These factors or features can be listed as:
System Feedback, Error Prevention, User Performance, Consistency, Error Recovery, Privacy, User Like / Dislike, Internationalization, Cognitive Load and Help.

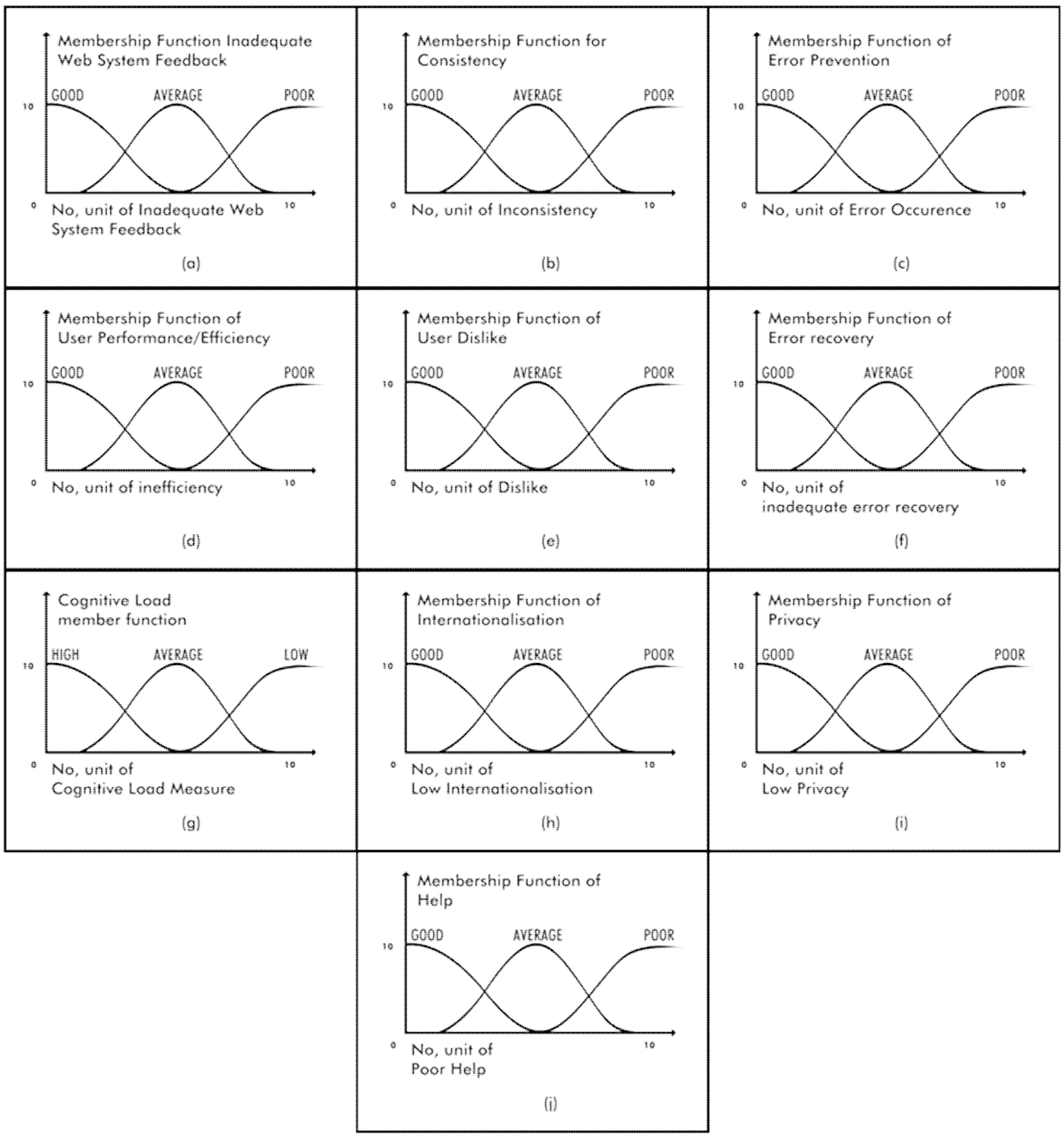
The input vector x is, therefore, defined to be
x= [System Feedback, Error Prevention, User Performance, Consistency, Error Recovery, Privacy, User Like / Dislike, Internationalization, Cognitive Load, Help]
We will write this using short form notation as

x = [FB, CO, EP, PF, UO, ER, CL, PR, IN, HE]

A typical premise would be of the form "Feedback is good".
The structure of the left hand side of the rule, therefore, is
IF ((WEB SYSTEM FEEDBACK is X1) AND (CONSISTENCY is X2) AND (ERROR PREVENTION is X3) AND (USER PERFORMANCE is X4) AND (USER LIKES/DISLIKES is X5) AND (ERROR RECOVERY is X6) AND (COGNITIVE LOAD is X7) AND (PRIVACY is X8) AND (INTERNATIONALISATION is X9) AND (HELP is X10)

X_i, i = 1,2,3,4,5,6,8,9,10 denotes in each case the fuzzy sets corresponding to the linguistic terms [Poor, Average, Good]. X7 denotes the fuzzy sets corresponding to the linguistic terms [Low, Average, High].

Figure 4.1: Member Functions



The actual form of the fuzzy sets and the associated membership function in each case were defined. In short form notation the left hand side of each of the rules would take the form:

IF ((FB is X1) AND (CO is X2) AND (EP is X3) AND (PF is X4) AND (UD is X5) AND (ER is X6) AND (CL is X7) AND (PR is X8) AND (IN is X9) AND (HE is X10)).

Since we are using the linear form of the first order Takagi Sugeno inference system the right-hand side for the rule q has the form

Web usability yq = aq0 + aq1 x1 + ... + aq10 x10 , where x1 , ..., x10 are the input variables

i.e. in short form notation –usability:

yq=aq0+aq1*FB+aq2*CO+aq3*EP+aq4*PF+aq5*UD+aq6*ER+aq7*CL+aq8*PR+aq9*IN+aq10*HE

Here aqo, aq1, ..., aq10, are parameters. A typical example of the left hand side in short form notation is

IF ((FB is GOOD) AND (CO is GOOD) AND (SN is GOOD) AND (EP is GOOD) AND

(DT is AVERAGE) AND (PF is AVERAGE) AND (UD is GOOD) AND (ER is GOOD) AND (CL is AVERAGE) AND (PR is AVERAGE) AND (IN is GOOD) AND (HE is POOR))

The form of the qth rule in the rule base, in short form notation, therefore, is

IF ((FB is X1) AND (CO is X2) AND (EP is X3) AND (PF is X4) AND (UD is X5) AND (ER is X6) AND (CL is X7) AND (PR is X8) AND (IN is X9) AND (HE is X10)) THEN aq0+aq1*FB+aq2*CO+aq3*EP+aq4*PF+aq5*UD+aq6*ER+aq7*CL+aq8*PR+aq9*IN+aq10*HE

The total number of possible fuzzy rules if we have n inputs and use K fuzzy sets to span the universe of discourse for each input is equal to K^n. For the usability problem, therefore, the total number of rules that we can have is 312 = 531441.

This takes into consideration all possible combinations of the inputs. This is a rather extensive number of rules and many of these rules are redundant and hence will be deleted, in order to make the system more manageable. Further we note, that for any input instance, 212 = the maximum number of rules that could be triggered.

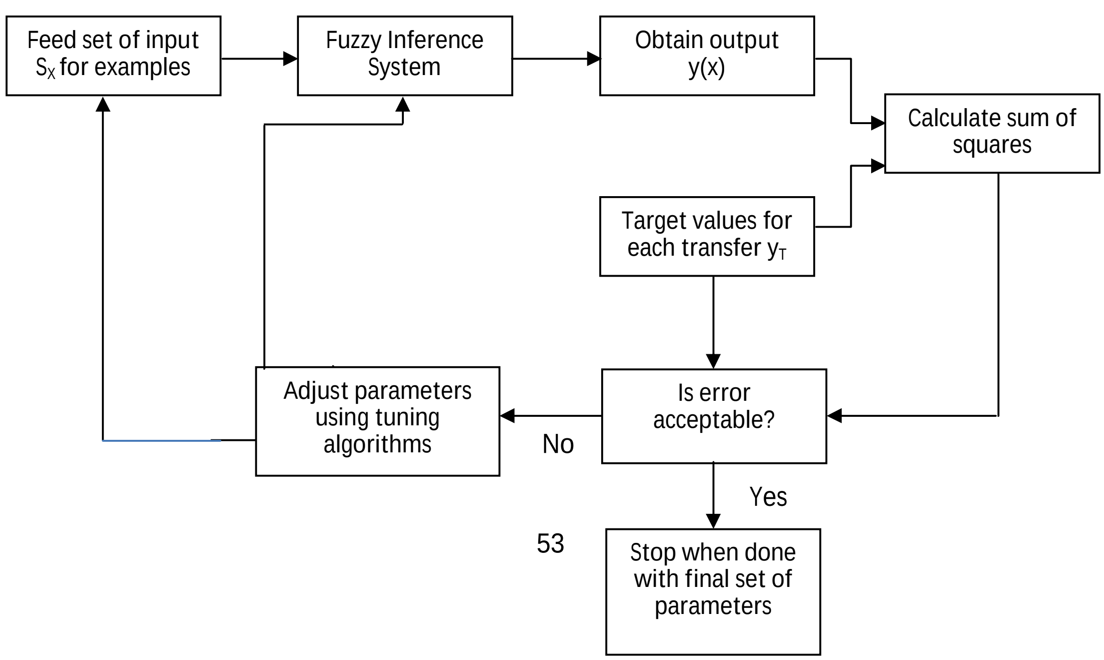
4.3.3 Tuning Algorithms For The Fuzzy Controller

As explained earlier, we need to tune the parameters that characterize the input membership function and the output crisp function associated with each rule. If we consider the most general form of the Takagi-Sugeno, (1985) model i.e. for rule q IF (x1 is X1) AND (x2 is X2) ...AND (xn is XN) THEN g(x1, ... , xn)

Then the problem of tuning the parameters is illustrated in Figure 2 below. The output associated with input example x^r is $y^r(x^r)$ and this is given for the linear Takagi-Sugeno controller by expression of:

$$y^r(x^r) = \frac{\sum_{q=1}^{M} \alpha q^r \left(\sum_{s=1}^{N} \alpha q^s \times s^r\right)}{\sum_{q=1}^{M} \alpha q^r} \quad \textbf{(4.4)}$$

Figure 4.2: Diagnostic Representation of Tuning Process



Here αq^r is the firing strength corresponding to rule q for the inputs x^r corresponding to the r. Let there be R examples altogether and let us write the target value of output corresponding to example r (or inputs x^r) as $y_T r$. Then the sum of the squares of the errors generated will be given by expression below.

$$E(\rho) = \left(\sum_{r=1}^{R} y_T r - y(x) \right)^2$$

(4.5)

The tuning algorithms that are available seek to minimize the value of this error function. There are many different tuning algorithms for doing these. We will only discuss the three main approaches to doing this and they are:

i. Least Square Error Optimization Approach
ii. Gradient Method
iii. Neural Net Method

4.3.4 Fuzzy Usability Control Rules

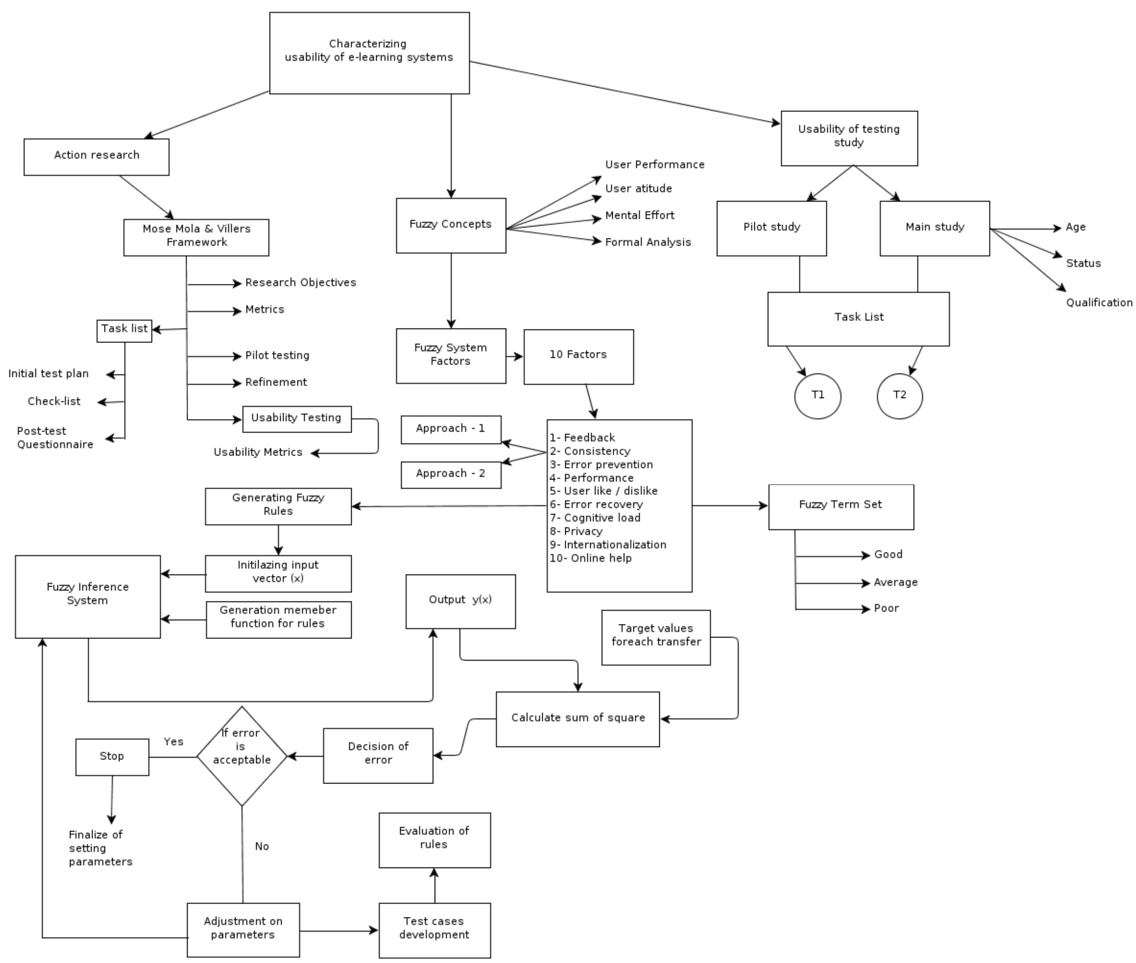
The individual fuzzy sets, i.e. the usability factors described above, will determine the result of the overall usability of the system being tested. Each usability value for a particular set of usability factor values will be determined through the use of Fuzzy Control Rules. If Feedback is good, Consistency is good, Error Prevention is poor. Error recovery is average then Usability is (some_value).

This is the rule in its simplest format. For the control rule to be intuitive there are a number of other factors which need to be considered. Each control rule has to be set up so that the output is some function of the inputs. The control rules are to be set up so that a linear relationship exists between the inputs and the outputs.

It is important to note that there are several control rules to properly model Usability. This is because a particular usability test may result in a particular set of usability factor values, and another particular instance may lead to a completely different set of parameter values. Both of these instances have to be modeled using separate control rules, in order to determine the corresponding value of the usability.

5. DATA COLLECTION AND ANALYSIS

Figure 5.1: Method Work Flow Diagram



In this section, a different approach for the usability is discussed in the context of e-learning.

5.1 USABILITY TESTING STUDY

For the module ELM (E-Learning Module used for the usability analysis) the participants were drawn from registered students in the first approach. As a requirement which is additional, those who were in the list of co-discovery testing have been asked to take part in a study group of the module. The participants who have been accepted for co-discovery testing may offer additional challenges; here the level of expertise of each of the pair can be necessary for consideration also as familiarity existence between these members.

In the pilot and main studies sixteen learners have been participated in total. As they were the representative of diverse learner population the participants were selected. For the pre-pilot test they were two participants, for pilot seven and the remaining seven were for the main test. The pilot study and the main study both included a set of co-discovery participants.

Five of the participants were males and two were females for the main test. While the other two study part-time, five were full-time students. All seven students were registered for a programming module. The profile of the main test participants are provided in the Table 5.1.

In order that the participants might not expose to physical danger by usability testing, together with the presence of recording cameras and sensation of being observed of the controlled nature of the test environment, can be a source of distress to some participants who can feel pressured to be well performed, also it was told them that the application being tested and not them.

Table 5.1: The main test participants profile (The* indicates a co-discovery participant)

No	Age	Computing experience	Status	Qualification
1	28	Average	Part-time student	Matric
2	31	Average	Full-time student	Diploma
3*	19	Minimal	Full-time student	Matric
4*	20	Minimal	Full-time student	Matric
5	22	Advanced	Part-time student	Matric
6	24	Minimal	Full-time student	Matric
7	19	Minimal	Full-time student	Matric
8*	21	Minimal	Full-time student	Matric
9*	21	Average	Full-time student	Matric

5.2 TASK LIST FOR UT

In this task list there are two tasks to be completed; please the tasks have to be completed in specified order.
Note: After you have completed each task please inform the test administrator.

Task 1:

i. Go to menu from the title page.
ii. Go to "Karnaugh diagrams" submenu from the main menu page.
iii. Your study in learning the presented content this section and doing the associated exercises is much appreciated.

Task 2:

i. Go to "A testing game" sub-menu from the main menu.
ii. Read and follow the instruction on how to select and answer questions.
iii. Make the actual test (Note: This version of the "testing game" is a subset which has been created specifically for this usability testing study).

6. FINDINGS

6.1 USABILITY TESTING STUDY FINDINGS

In total of nine participants' sample is very small for statistical analysis; that's why qualitative interpretations were mainly made on the co-discovery wise. It is a vital requirement for measuring use-of-time patterns i.e. recording the time spent in navigation and reading screen information and the time spent on actual learning activities using by participants of the think-aloud method. The two sets of co-discovery test participants found thinking aloud and the discussion during the sessions of their activities to be a comfortable process. This was expected, because it emerged naturally as two people conversation.

Contrary, the participants of the single test were silent during the big part of the sessions; they found it difficult to think aloud, even there was prior coaching on the process. Several authors, including and aware of this problem that arises with single participant think aloud.

We observed during the testing sessions many learning styles and group dynamics, that, according to our opinion, can contribute to collaboration and pear-teaching among learners. The observations were included the followings:

i. In the group A's case: GAP 1 was in charge of the computer for the two tasks, by making all navigational decisions requested and to type the answers to exercises. For Group B, GBP1 and GBP2 have swapped the roles, with GBP1 which takes charge for task 1 and GBP2 which takes charge for task 2.

ii. Each of the groups has negotiated their progression through the lessons, with the participant who controls the computer and who asks the partner if they could progress.

iii. In group A, GAP1 became to be the “stronger” student; anyway, this participant did not move on till the “weaker” one (GAP2) had comprehended in full the given concept. This is reflected in, that shows the patterns of use-of-time for group A since they complete the activities in task 1.

iv. The participants to GAP1 have used 16 percent of the time in total by explaining learning concepts to her partner. This is contrary to group B, where 8 percent of the time was spent by G BP2 explaining some concepts to GBP1.

v. The patterns of time distribution between two groups are fairly similar, along with the participants in group A who have used 53 percent of their time in total by reading, learning related content and instructions, while those others in group B have used 60 percent of their time in total for the same activities. The two way discussion of the concepts that have been consumed as 11 percent and 13 percent of time respectively; while 14 percent and 11 percent was taken up with the answers to the section exercises.

vi. Respectively two-way discussion of concepts of time consumed 11 percent and 13 percent respectively; while 14 percent and 11 percent was taken up by answering section exercises.

vii. The two groups used a mouse combination and manual touching of the screen, where the participant who did not handle the mouse did the hand-touching.

viii. The two used paper and pencil for rough work and calculations.

ix. The two groups before calling up the system-provided answers have worked out their own solutions to segment examples in task 1.

x. Each of the participants in group B for the exercises in task 2, worked on the exercises independently, then before typing or selecting their answers have compared them. Anyway, in the case of group A, the ‘stronger’ participant (GAP1) did most of the work, as well as taking time to explain reasoning behind the answers to her partner.

The difference of major time is the only common question (Q3*) answered by both groups, where the group A pair spent 564.2 seconds on Q3 while their counterparts in group B spent 298.2 seconds on the same question.

Table 6.1: The metric performance of co-participant for task 1.

Working through the content of Karnaugh diagrams, including section exercise (time shown in seconds)		
Activity	**Group A**	**Group B**
Number of mouse clicks	44	42
Number of calls for help	0	0
Number of usability errors	0	0
Number of correct answers	3	3
Navigation time	20	25.9
Time spent reading	673.4	952.6
Time spent on discussion	144.6	206
Time spent explaining concepts	211.2	130.8
Time spent working on examples	55	89
Time spent on section exercises	174.4	178.6
Time spent learning/peer teaching	585.2	604.4
Total completion time	1278.6	1582.9

Figure 6.1: Group A: Time Distribution, Task 2

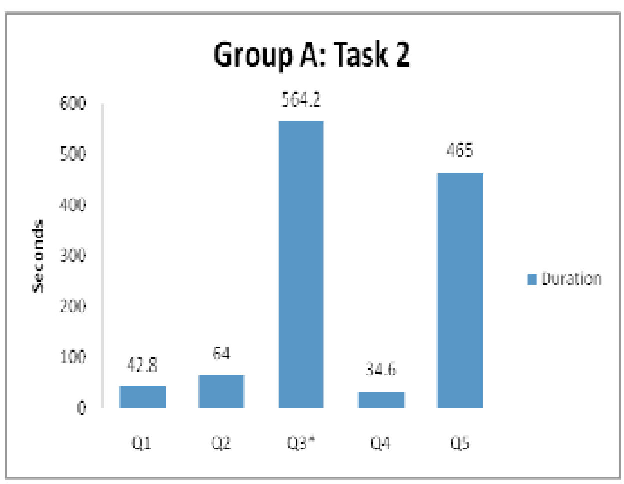
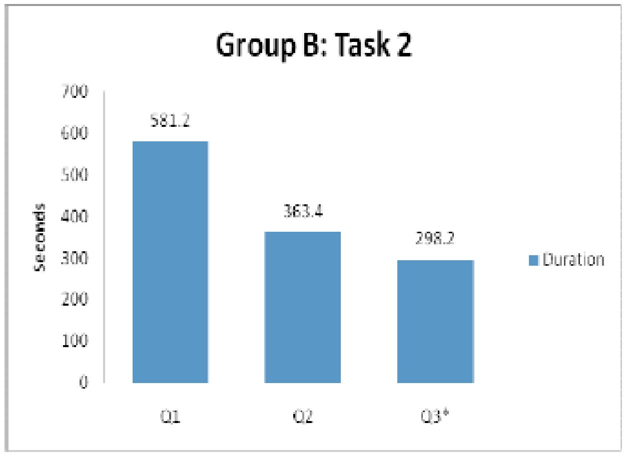


Figure 6.2: Group B: The Distribution of time for Task 2



As shown in Table 6.2, the participants of group A answered 5 questions and used 1170.6 seconds in total to complete them while the peer in group B answered 3 questions, spending 1242.8 seconds on the questions. It is notable that the two groups gave correct answers to the questions on their first attempts.

Because of the difficulty experienced by the single test participants in thinking aloud during the testing sessions, we were not able to distinguish between time spent on reading and navigating, and time spent actually learning concepts. Tables 6.3 and 6.4 show the metric performances for single participants for activities in tasks 1 and 2 respectively. From all five participants, only one (participant 6) did not require some form of content-related help from the test administrator (who was the primary researcher) while completing task 1.

Table 6.2: The metric performance of Co-participant for task 2

Working through exercises in task 2 (time in seconds)			
Group	Number of questions answered	Number of correct answers	Total time spent on all questions
A	5	5	1170.6
B	3	3	1242.8

In every case, time spent on help was deducted from the total completion time. Participants 2 and 5 have been struggled most with content comprehension, which resulted inconsiderably more time spent on task 1 in comparison to the others. In task 2, the number of questions ranged from four to ten answered by the participants, with four common questions answered by all. In a similar pattern to task 1, participants 2 and 5 spent the most time on task 2.

Table 6.3: The Single metric performances of participant for task 1

Working through the content of Karnaugh diagrams, including section exercise (time shown in second)					
Activity/ Participant	**1**	**2**	**5**	**6**	**7**
Number of mouse clicks	44	85	50	43	58
Number of calls for help	2	2	1	0	1
Number of usability errors	0	0	0	0	0
Navigation time	40.6	70.4	38.6	37.1	48.2
Intervention time by test administrator	25.2	26.6	29.4	0	139.6
Time spent on section exercises	62.2	952.6	270.2	89	437.2
Total completion time	426	3593.4	2437.2	813	1430

Table 6.4: The Single metric performances of participant for task 2

Working through exercises in task 2 (time shown in second)					
Activity/ Participant	**1**	**2**	**5**	**6**	**7**
Number of usability errors	1	2	1	0	1
Number of questions answered	9	10	7	4	5
Number of correct answers	5	2	4	4	4
Total time spent on questions	804.2	2655.5	3378.4	1018.4	1223.2

Among the most notable findings one was the natural way in which the co-participants were able to think aloud, when compared to single participants. It was of interest too that, while the single participants required content-specific assistance from the test administrator, in the two sets of co- participants, this type of help was provided by one of the participants.

One of the main advantages of performing co-discovery testing is providing learning environment without interacting with the teacher and collaborative learning measure is suitable for e-learning platform.

The other interesting observation was the time spent completing the activities in tasks 1 and 2. As shown in Table 6.1, the co-participants in groups A and B spent 1278.6 and 1582.9 seconds respectively in total on task 1. Even, when times spent on discussions with partners and explanations of concepts are deducted, accounting for 355.8 and 336.8 seconds respectively, the times in total are reduced to 922.8 seconds for group A and 1246.1 seconds for group B.

Comparison of time spent on tasks by single and co-participants indicates that co-participant testing does not necessarily increase the duration of time. For instance, as shown in Table 6.5, only the participants 1 and 6 have spent less time on task 1 than the co-participants, and participants 2 and 5 have spent twice as much time as the co-participants on the same task! In fact, the co-participant testing may even results in less time as participants assist each other through the learning content.

It should even be noted due to the small number of co-participants involved in the study that no general conclusions can be drawn from these observations. The participants who have used paper and pencil for rough work performed well both single participants and co-participants in the exercises.

To each participant the post-test questionnaires comprising semi-structured questions were given to measure their subjective perception of Karnaugh. The participants who performed relatively well viewed the structure of Karnaugh as ‘very easy’ or ‘easy’ while those who did less well found it ‘very difficult’ or ‘difficult’. All the participants liked the use of colors to emphasize important information.

Two students found the different representations of Karnaugh diagrams to be confusing (the tutorial provides four different ways of representing Karnaugh diagrams while the printed study guide provides only one representation). All the participants viewed the tutorial as a useful learning tool and would like to have similar tutorials for other modules.

Table 6.5: Comparison of the spent time on tasks 1 and 2 by participants

Duration of tasks 1 and 2 (time shown in seconds)		
Participants	**Time on task 1**	**Time on task 2**
Group A	922.8 (excluding discussions and explanations)	1170.6
Group B	1246.1 (excluding discussions and explanations)	1242.8
1	426	804.2
2	3593.4	2655.5
5	2437.2	3378.4
6	813	1018.4
7	1430	1223.2
Average time	1552.6	1641.9

As well as Karnaugh had been evaluated previously by using various UEMs such as heuristic evaluation, questionnaire surveys and interviews, formal UT involving real users doing real tasks showed some previously unidentified usability problems that could distract from learning. The application that illustrates the monitor view seen by the test administrator during UT is a simultaneous triple screen display. The displays, clockwise from left are:

i. current screenshot of the application being evaluated,
ii. the keyboard – at the instant being recorded, the participant's hands are not on the keys,
iii. the participant's facial expression with his hand near his face. Displays such as this during the evaluation provide a holistic view to the test administrator. This subsection highlights some of the problems identified and makes recommendations for improvement.

a. The meaning and use of the multi-option control button, offering the composite functionality of >I<, >E<, >QB<, and >QC<, representing direct access to information, worked-out examples, selection of a basic question, and a complex question, respectively, are not intuitive. Although increased learner control was provided by this useful feature and was rated highly in an earlier questionnaire survey, none of the seven participants in the main test used it. This is possibly because of the insufficient in-context information regarding its functionality.

Even an introductory section of Karnaugh provides the meanings of the button in a decontextualized way; the learners have no easy access to this once that they have commenced the lesson. The designers cannot expect that the learners have to remember the meanings throughout their interactions with an application. We advised the designer to use roll-overs to emphasize the meanings within the button.

b. A participant was using a screen that offers phased development of a concept. Each click on the >More< button provided further information. He finally wanted to review the content of that screen and clicked the >Back< button, expecting to return to the beginning of the progression and watch the concept development again. He was, even, returned to the last exercise of the previous segment. The tutorial designer had provided a >Repeat< button for his required purpose which he overlooked, since it created a mismatch between the designer and learner’s model, which is in violation of a Squires and Preece heuristic.

Additionally, the >Repeat< button was grayed out on the screen. We recommended that adequate information should be provided on the intended use of each button in clear and unambiguous terms. The >Repeat< button, which had been erroneously grayed out, should be re-activated to make it accessible.

c. In the ‘Testing game’ section, in one of the exercises, the four participants have been provided answers that were longer than the space provided in an answer text box. To check their answers, the participants had to use the arrow key on the keyboard because the incomplete answers were not visible.

While we acknowledge that the short space provided for learner response is often an important clue for the required length of the answer, we recommended that the number of allowable characters should be fixed as a form of forcing function, thus improving the visibility of the answer.

In the second approach called the fuzzy system, reliable measures for usability for input vectors are produced that are close to the values of the training input-output pairs for each of the interfaces used in the testing. The fuzzy system is able to intuitively predict the direction the overall usability should go when particular inputs are increased and decreased relative to a particular training data vector.

When feedback, consistency, error prevention, performance, and error recovery are increased, the overall usability value generally decreases, which is what is intuitively expected. Also if these factors are decreased, then the overall usability is increased which is depicted in the above tables.

These results depict the importance that feedback and consistency have on the overall usability. When these two factors are increased and decreased, the usability significantly changes in the expected direction. In Table 5.1 when the feedback value in vector is changed from 6 to 3 the usability increases from 6.5000 to 9.0028.

When the consistency value of test 4 is increased from 2 to 6, the usability decreases considerably from 8.3482 to 3.5336. This test shows the importance that feedback and consistency have in the FIS at affecting the overall usability, which is representative of real-world results.

The usability factor of opinion also behaves in a consistent manner. Increasing and decreasing the user opinion affects the usability value in the same direction. These tables highlight the significant effect that opinion has on the overall usability.

A reasonable increase or decrease in opinion changes the value of the overall usability considerably. In Table 6.2, has a usability value of 6.9287, when the user opinion is changed
from 6 to 10 the usability increases significantly to 9.1340. This is what is expected from the features of the training data. Of all the inputs, opinion is the one which resembles the desired usability rating across all training pairs. Hence, significantly altering the user rating will cause a similar change in the output value for usability.

6.2 FUZZY LOGIC STUDY FINDINGS

6.2.1 Output Function for Usability

As the fuzzy inference method used is the T-S model the output membership functions are singleton spikes, rather than fuzzy sets. Each count of the parameters FB, CO, EP, PF, UD, ER, CL, PR, IN, HE, constitutes a deficiency or inadequacy towards the corresponding usability measure. Hence the higher the count of these web usability measures, then the worse it is (the more likely it is to be situated in the 'Poor' region) which would indicate a poor usability.

However, the Web Usability output generated is the reverse of this, since the higher value of generated output the more optimum the web usability is. This is to keep the results of the Fuzzy Inference System intuitive, since one would expect a high value of usability for an interface to represent a good e-Learning system.

The usability output generated will be a continuous value between 0-10, and so a value of zero would represent an extremely poor interface, whereas a value of ten would indicate that the UI has an outstanding usability.

6.2.2 Evaluation of The Fuzzy Rules

The parameters of the fuzzy rules are generated through examination of the local features of the training data. Hence the fuzzy control rules would seem to produce a reliable measure for input cases that are exact or close to the training data vectors. Input values not in the original training set were run with the set of fuzzy rules and it was found that the fuzzy rules predicted the results well.

The following results show the usability values obtained by plugging in varying types of input vectors into the fuzzy system. The order of the usability parameters of the input vectors are [feedback, consistency, error prevention, performance, error recovery, user opinion].
Input vector in the tables are indicated as: [FB, CO, EP, UP, PE, ED, UD, ER, HE].

Table 6.6 shows the input test cases derived from varying two different input-output training pairs obtained from testing the 1st Good interface. These input vectors are created by individually increasing and decreasing the value of each usability factor of the original training data vector.

Table 6.7 shows the input test cases derived from varying two different input-output training pairs obtained from testing the Average interface.

Table 6.8 shows the input test cases derived from varying two different input-output training pairs obtained from testing the Poor interface.

Table 6.6: TC for "Good" ***Table 6.7: TC for "Avg"*** ***Table 6.8: TC for "Poor"***

Test Case	**Input Vector**	**Usability**	**Test Case**	**Input Vector**	**Usability**	**Test Case**	**Input Vector**	**Usability**
1	[0,0,2,0,1,1,1,0]	9.3625	**1**	[4,6,2,3,1,3,1,4]	6.8594	**1**	[8,9,4,7,8,7,6,7]	2.7513
2	[1,0,2,0,1,1,1,0]	9.2780	**2**	[5,3,5,2,1,3,2,3]	6.8755	**2**	[8,9,6,7,8,7,6,7]	2.4452
3	[0,1,2,2,0,3,2,0]	8.6595	**3**	[4,3,5,2,2,3,3,3]	7.0625	**3**	[4,6,10,7,10,5,8,7]	2.9914
4	[3,1,1,1,0,0,1,2]	8.6847	**4**	[6,2,0,7,4,4,8,2]	5.3373	**4**	[4,4,10,7,10,5,8,7]	2.9908
5	[2,2,3,0,0,2,2,1]	8.2833	**5**	[6,4,0,7,4,4,8,2]	5.3379	**5**	[8,9,8,7,8,8,6,8]	2.0148
6	[0,0,2,3,2,1,3,2]	8.2870	**6**	[8,9,9,7,8,8,6,8]	5.0759	**6**	[8,9,9,7,8,8,6,8]	1.9567
7	[1,4,1,1,2,2,1,2]	8.2975	**7**	[1,4,6,2,6,1,6,5]	6.0988	**7**	[2,9,10,8,9,8,6,8]	2.6624
8	[1,2,0,2,4,4,3,2]	7.7440	**8**	[1,4,5,2,8,2,6,2]	6.7517	**8**	[5,7,5,9,8,10,10,7]	2.7737
9	[3,3,4,2,3,3,2,1]	7.6887	**9**	[0,5,4,5,3,2,3,3]	6.6173	**9**	[5,5,5,9,9,10,10,6]	2.6156
10	[1,4,0,2,4,4,3,2]	7.7779	**10**	[4,3,5,2,2,4,3,5]	6.6264	**10**	[10,6,10,4,4,9,9,8]	2.6472
11	[4,3,3,2,6,4,3,1]	7.0687	**11**	[2,2,1,4,3,6,5,3]	5.2078	**11**	[10,10,6,10,8,6,7,6]	2.1831
12	[2,1,5,4,4,1,3,1]	7.0244	**12**	[6,6,4,5,2,1,7,3]	5.0684	**12**	[6,7,10,10,10,6,7,8]	1.9795
13	[0,1,1,2,2,0,0,2]	9.0647	**13**	[6,6,5,5,2,1,7,3]	5.2720	**13**	[6,10,7,7,7,9,10,10]	1.7256
14	[1,2,2,0,0,1,1,2]	8.9672	**14**	[6,6,7,1,2,1,8,3]	5.2414	**14**	[10,9,10,9,9,3,4,10]	1.6987
15	[0,2,1,1,0,0,2,1]	8.8174	**15**	[6,6,9,1,2,1,8,3]	6.4825	**15**	[9,6,10,8,8,10,10,6]	1.7738

7. DISCUSSION

The uniqueness of e-learning proposed by several researchers and various approaches are conducted for evaluating. For example, Squires et al. (1999), developed a set of criteria, called 'learning with software heuristic' that are based on Nielsen's heuristics, (1995) and the socio-constructivist learning paradigm. This set of heuristics provides educators with an effective tool for assessing the quality of educational software both for interface usability and pedagogical effectiveness.

With regard to UT steps, tasks and procedures, Masemola et al. (2006) proposed a framework for UT of e-learning applications. Their study involved formal testing in a controlled laboratory environment, with learners carrying out specified tasks using an e-learning tutorial called Relations. Valuable metrics in the form of time usage pattern allowed distinction to be made between time spent navigating and time spent on learning activities. Data analysis provided insight into differences in learning styles.

As an extension of the Masemola and de Villiers UT framework, (2006) for e-learning, this study applies and refines it using both single user testing and co-discovery testing.

The second study, fuzzy system produces reliable measures for usability for input vectors that are close to the values of the training input-output pairs for each of the interfaces used in the testing. The fuzzy system is able to intuitively predict the direction the overall usability should go when particular inputs are increased and decreased relative to a particular training data vector.

When feedback, consistency, error prevention, performance, and error recovery are increased, the overall usability value generally decreases, which is what is intuitively expected. Also if these factors are decreased, then the overall usability is increased which is depicted in the above tables.

These results depict the importance that feedback and consistency have on the overall usability. When these two factors are increased and decreased, the usability significantly changes in the expected direction. In Table 6.7 when the feedback value in vector is changed from 6 to 3 the usability increases from 6.5000 to 9.0028.When the consistency value of test 4 is increased from 2 to 6, the usability decreases considerably from 8.3482 to 3.5336. This test shows the importance that feedback and consistency have in the FIS at affecting the overall usability, which is representative of real-world results.

The usability factor of opinion also behaves in a consistent manner. Increasing and decreasing the user opinion affects the usability value in the same direction. These tables highlight the significant effect that opinion has on the overall usability. A reasonable increase or decrease in opinion changes the value of the overall usability considerably.

In Table 6.8**,** has a usability value of 6.9287, when the user opinion is changed from 6 to 10 the usability increases significantly to 9.1340. This is what is expected from the features of the training data. Of all the inputs, opinion is the one which resembles the desired usability rating across all training pairs. Hence, significantly altering the user rating will cause a similar change in the output value for usability.

Existing approaches for evaluating the usability of a system which are empirical testing, inspection, comparative usability measures, and formal complexity based measures have drawbacks. Empirical testing entails conducting usability evaluation in a fairly controlled situation to identify the problems that the user have with the system. Inspection involves examining the design of various screens to identify problems that might occur in the user interfaces.

Comparative usability measures consists of usability metrics that are designed to evaluate aspects of the interface such as the ability of the user to perform the tasks that s/he is required to do efficiently, and acceptance by the user of the user interface. Finally, formal complexity based measures on the other hand use a variety of criteria that rely on some formal or mathematical analysis of the user interface. With this fuzzy logic study the development of a comparative usability measurement approach for e-learning systems is conducted and three components are listed as heuristic evaluation, testing aspect, a usability measure that consists of the results of the tests. The framework we employ for developing the usability measure is Fuzzy Systems.

8. CONCLUSION

A useful and effective working structure for the study has been provided by the framework. Two pilot studies instead of one have been conducted in order to adapt it to the situation, by this way the necessary expertise could be acquired by the test administrator. Two sets of pairs were involved in the testing process as co-participants as addition to the testing of single participants. During the pair interactions some positive group dynamics were observed.

One of the most important aims of this study was to investigate the ease of the single and co- participants in UT could think aloud with. Thinking aloud is an essential requirement for distinguishing between time spent reading-and-navigating and time spent studying/processing content.

The participants of the single test have been struggled with think-aloud and mainly remained silent during testing sessions, making it difficult to make such distinctions. Anyway, for the co-participants, thinking aloud naturally came because it involved two people having a conversation. Additionally, the co-participant testing does not necessarily result in increased amount of time spent on activities.

It has been revealed by the two observations of co-discovery that the co-participant testing has the potential to reduce the level of intervention by the test administrator, and is especially relevant for testing e-learning, where collaboration is currently promoted as a useful form of learning. Even the number of pairs is too low for reliable generalizations the two cases demonstrated that it is possible for learners to assume the role of peer-teacher.

The collaborative learning and peer-teaching could play valuable roles when extrapolating beyond the immediate context of usability testing to the situation of learners in a distance-teaching context such as UNISA. The number which is limited of participants who have been involved in the study did not allow for statistical analysis of learner performance. More empirical research involving a greater number of participants, both single- and co-participants could be undertaken to validate the findings of this study.

The research reported in this paper can be described as 'meta-evaluation' in that it is focused not primarily on the target system, but rather on investigating two UEMs, in particular, heuristic evaluation, studying the extent to which HE by a small set of appropriate experts can identify usability problems in a web-based learning application. This was done by comparing the HE results with those of an end-user survey among the learner population, using evaluation criteria custom-generated for e-learning applications.

The findings of this study indicate that heuristic evaluation, if conducted by a competent and complementary group of experts, is an appropriate, efficient, and highly effective usability evaluation method in the context of e-learning, as well as relatively easy to conduct and inexpensive.

The researchers recommend that HE should, ideally, be supplemented with methods where users themselves identify usability or learning problems. This is in line with proposals that with user-based inspection methods being combining systematically the reliable evaluation can be achieved. In cases where only one approach has to be selected, the findings of this research can be used to propose heuristic evaluation as the optimal method.

A usability measurement approach for e-learning was discussed in this paper. We utilized some traditional usability measures and recent results based Usability Concepts. Data gathered on the actual e-learning system using Automated Usability Tool was used to tune the parameters of the Fuzzy Model. The results obtained from the Automated Usability Software indicate that it is a suitable measure to e-learning.

The summarized results of usability assessment, along with the users' comments and observations collected upon performing evaluation sessions, indicate the necessity of a number of improvements in each and every design. Both teachers and students found the arrangement of domain knowledge somehow confusing, thus preventing them to quickly understand the domain knowledge structure.

In order to hide as much as possible the internal structure of the domain knowledge base, the knowledge presentation was redesigned in both the on-site and the web-based e-learning system. This will hopefully result in: more transparent and intuitive, faster creation of a specific knowledge base, and faster learning of a specific subject matter.

It is possible to enrich the e-learning experience via a well-designated interface – The design of the learning interface enriches the e-learning experience, but up to a certain level. In nearly any parameter, the learner can realize his learning potential via use of a well-designed interface, but may not reach full abilities if the interface is not well designed.

An interface wrongly designed can cause a decline in the perception of self-adaptation – When the system is designed in a way that doesn't support the needs of the learner and does not provide a qualitative experience, the learner might, despite a good starting point, feel that he is losing his abilities, leading to a decline in his perception that he can succeed.

The flow experience of the new learner turns out to be innovative. Despite the claim of Salomon that the depth of learning is hurt by the butterfly defect, it seems that it could be fixed by an interface well designed. The learner keeps distributing his attention between several parallel activities, but simultaneously, as the e-learning system supports these abilities – he has many flow states without disruption to his line of thought.

In general, the new learner does find the process of distributed learning and the passing of responsibility important as an empowering process. But as a old learner with a conservative thinking style finds himself passing the responsibility for learning to someone else, instead of just clinging to the teacher, he finds an additional source to be assisted by – new learners that have been empowered and have become opinion leaders and tone setters in the group learning process.

A usability measurement approach for e-learning was discussed in this thesis. We utilized some traditional usability measures and recent results based Usability Concepts Data gathered on the actual e-learning system using Automated Usability Tool was used to tune the parameters of the Fuzzy Model. The results obtained from the Automated Usability Software indicate that it is a suitable measure to e-learning.

Overall, we have found that a richer user experience will enable most learners to realize a more qualitative e-learning experience. However, further, deeper investigation and expanded experimentation are still merited.

REFERENCES

Books

Norman, D., 1993. *Things That Make us Smart: Defending Human Attributes in the Age of the Machine*, Perseus Publishing.

Rossett, A., 2002. *The ASTD ELearning Handbook: Best Practices, Strategies, and Case Studies for an Emerging Field*, McGraw Hill.

Notess, M., 2001. *Usability, User Experience, and Learner Experience*.

Shneidermann, B., 1997. *Designing the User Interface*, Addison-Wesley.

Nielsen, J., 1993. *Usability Engineering*, Boston, MA: Academic Press.

Ravden, S.J. Johnson, G.I., 1989. *Evaluating Usability of Human-Computer Interface: a Practical Method*, John Wiley & Sons.

Nielsen, J. Mack, R.L.,1994. *Usability Inspection Methods*, John Wiley & Sons, pp: 5 - 12.

Nielsen, J., 1994. *Heuristic Evaluations, Usability Inspection Methods*, John Wiley & Sons.

Nielsen, J. 2000. Designing Web Usability.

Periodicals

Takagi, T. Sugeno, M., 1985. Fuzzy identification of systems and its applications to modelling and control, *IEEE Trans. Syst., Man, Cybern.*, **15**, (1), pp 116-131.

Charles, A. Shoniregun, Gray, S.J., 2003. Is E-learning really the future or a risk?, *ACM Ubiquity.*

Karl M.K., 2003. E-learning basics: essay: Five technological considerations when choosing an e-learning solution, *ACM eLearn.*

Feldstein, M., 2002. E-learning basics: essay: developing your e-learning for your learners, *ACM eLearn.*

Molich, R. Nielsen, J., 1990. Heuristic Evaluation of User Interfaces, *Proceedings of ACM Human Factors in Computing Systems,* pp: 249–256.

Nielsen, J. Philips, V., 1993. Estimating the relative Usability of Two Interfaces: Heuristic, Formal, and Empirical Methods Compared. *ACM/IFIP Human Factors in Computing Systems*, pp: 214 - 221.

Dillon, T. Maquire, M., 1993. Usability Measurement- Its Practical Value to the Computer Industry. *ACM/IFIP Human Factors in Computer System*, pp: 145 - 148.

Quintana, C. Carra, A. Krajcik, J. Elliot, S., 2001. Learner-Centred Design: Reflections and new Directions, Human-Computer Interaction in the New Millennium, in Carroll (Ed.) ACM Press, pp: 605 - 626.

Squires, D. Preece, J., 1999. Predicting Quality in Educational Software: Evaluating for Learning, Usability, and the Synergy between them, *Interacting with Computers, Elsevier Science,* **11**, (5), pp: 467 - 483.

Kukulska-Hulme, A. Shield, L.,2004. Usability and Pedagogical Design: are Language Learning Websites Special?, AACE Digital Library, pp: 4235 - 4242.

Squire, D. Preece, J., 1996. Usability and Learning: Evaluating the potential of educational software, Computer & Education, Elsevier, **27**, (1),pp: 15 - 22.

Quinn, C. Alem, L. Eklund, J., 1999. A pragmatic evaluation methodology for an assessment of learning effectiveness in instructional systems, INTERACT99, **2**, British Computer Society, pp: 55 - 56.

De Marsico, M. Levialdi, S., 2004. Evaluating web sites: exploiting user's expectations, International Journal of Human-Computer Studies, **60**, pp: 381 - 416.

Noiwan, J. Norcio, A.F., 2006. Cultural difference on attention and perceived usability: investigating color combinations of animated graphics, International Journal of Human-Computer Studies, **64**, pp: 103 - 122.

Beaudoin, M., 2002. Learning or lurking? Tracking the "invisible" online student, The Internet and Higher Education, **5**, pp: 147 - 155.

Squires, D., 1999. Usability and Educational Software Design: Special Issue of Interacting with Computers, Interacting with Computers, Elsevier Science Ltd, **11**, pp: 463 - 466.

Jeffriess, R. Miller, J. Wharton, C. Uyeda, K.M., 1991. User Interface Evaluation in the Real World: a Comparison of Four Techniques, Proceedings of CHI'91, ACM Press, pp: 119 - 124.

Matera, M. Costabile, M.F. Garzotto, F. Paolini, P., 2002. SUE Inspection: an Effective Method for Systematic Usability Evaluation of Hypermedia, IEEE Transactions on Systems, Man and Cybernetics- Part A, **32**, (1), pp: 93 - 103.

Blandford, A Keith, S. Connell, I. Edwards, H., 2004. Analytical Usability Evaluation for Digital Libraries: A Case Study, ACM/IEEE-CS Joint Conference on Digital Libraries, ACM Press.

Dubey, S.K Mittal, A. Rana, A., 2012. Measurement of Object Oriented Software Usability using Fuzzy AHP, *International Journal of Computer Science and Telecommunications*, **3**.

Other Sources

Masemola, S.S. de Villiers, M.R., 2006. Towards a Framework for Usability Testing of Interactive E-Learning Applications in Cognitive Domains, *Annual research conference of the South African Institute of Computer Scientists and Information Technologists (SAICSIT)*, pp. 187-197.

Teo, C.B. Gay, R.K.L., 2006. A knowledge-driven model to personalize e-learning, J. Educ. Resour. Comput. **6**, (1), Article 3.

Lytras, M., 2002. E-learning pedagogy: The reveal of value adding learning processes. Definitions and Implications for dynamic learning content delivery. In M. Driscoll & T. Reeves (Eds.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*, pp: 1826-1830.

Wong, S.K. Chang, E., 2002. Usability Measure For Web Based Applications, *Software Metrics Conference*, Melbourne, Australia.

Polsani, P.R., 2003. Use and abuse of reusable learning objects, JoDI – J. Digital Inf.3, **4**.

Doroodchi, M. Nikmehr, N., 2008. Overview and Evaluation of Usability Evaluation Methods for e-Learning Systems, *Proceedings of the 2008 international conference on e-Learning, e-Business*, Enterprise Information Systems, and e-Government. LasVegas: EEE'08.

Fetaji, B. Pop-Jordanova, N. J. Zorcec, T. Markosvka, S., 2007. Measuring e-Learning effectiveness through e-Content and attention correlation, *Proceedings of the 2nd International conference on virtual learning*.

ISO. 1998. ISO 9241: Guidance on Usability Standards.

Shackel, B., 1981. The concept of usability, *Proceedings of IBM Software and Information Usability Symposium*, pp: 1–30.

Ssemugabi, S. de Villiers, R., 2007. A comparative study of two usability evaluation methods using a web-based e-learning application, *In Proceedings of the 2007 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries*, pp: 132-142.

Ruttenbur, B. Spickler, G. Lurie, S., 2000. E-Learning: The Engine of the Knowledge Economy, *Morgan Keegan & Co. Inc*.

de Villiers, M.R., 2005. e-Learning artefacts: are they based on learning theory?, pp: 345-371.

Trace, U. Cornelia, W., 2000. Corporate e-learning: Exploring a new frontier, Equity Research, WR Hambrecht+Co.

Thimbleby, H., 1994. Formulating Usability, *SIGCHI Bulletin*, pp: 59 - 64.

Bavan, N. Macleod, M., 1994. Usability measurement in context. Behaviour and Information Technology, pp: 132 - 145.

Chang, E. Dillon, T.S. Cook, D., 1996. Measurement of Usability of Software Using a Fuzzy System Approach, *Invited paper Proceeding of Eighth International Conference on Software Engineering and Knowledge Engineering*, pp: 69 - 76.

ISO, 1998. ISO 9241: Ergonomics Requirements for Office Work with Visual Display Terminal (VDT).

Cha, H.J. Kim, Y.S. Park, S.H. Yoon, T.B. Jung, Y.M. Lee, H., 2006. Learning styles diagnosis based on user interface behaviors for customization of learning interfaces in an intelligent tutoring system, *ITS2006*, pp: 513 - 524.

Baker, K. Greenberg, S. Gutwin, C.,2002. Empirical Development of a Heuristics Evaluation Methodology for Shared Workspace Groupware, *Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW'02)*, pp: 96 - 105.

Mankoff, J. Dey, A. Hsieh, G. Kients, J. Lederer, S. Ames, M., 2003. Heuristic Evaluation of Ambient Display, *Human Factors and Computing Systems*, pp: 169 - 176.

Somervell, J. Wahid, S. McCrickard, D.S., 2003. Usability Heuristics for Large Screen Information Exhibits, Human-*Computer Interaction (Interact '03),* pp: 904 - 907.

Schwier, R.A. Misanchunk, E.R., 1993. Interactive Multimedia Instruction, NJ: Educational Technology Publications.

Zaharias, P., 2005. E-learning design quality: A holistic conceptual framework, *Encyclopedia of Distance Learning*, 2.

Nielsen, J., 1992. Finding Usability Problems through Heuristic Evaluation, *Proceedings of the SIGCHI Conference on Human factors in Computing Systems*, pp: 373 - 380.

Belkhiter, N. Boulet, M. Baffoun S. Dupuis, C., 2003. Usability Inspection of the ECONOF System's User Interface Visualization Component, C. Ghaoui. (Ed.), *Usability Evaluation of Online Learning Programs, Information Science Publishing*.

Hosie, P. Schibeci, R., 2001. Evaluating Courseware: A need for more context bound evaluations?, *Australian Educational Computing*, pp: 18 - 26.

Benson, L. Elliott, D. Grant, M. Holschuh, D. Kim, B. Kim, H. Lauber, E. Loh, S. Reeves, T., 2002. Usability and Instructional Design Heuristics for E-Learning Evaluation, *Proceedings of World Conference on Educational Multimedia,Hypermedia and Telecommunications*, pp: 1615 - 1621.

Ardito, C. Costabile, M.F. De Marsico, M. Lanzilotti, R., Levialdi, S. Roselli, T. Rossano, V., 2006. An Approach to Usability Evaluation of E-Learning Applications, *Universal Access in the Information society*, pp: 270 - 283.

Yahaya, Y.H. Mohamad, N., 2011. Designing Software Usability Measurement Using Fuzzy Set Conjoint Model*, International Conference on Computer Communication and Management,* **5**.

Liang, L. Deng, X. Wang, Y., 2009. Usability Measurement Using a Fuzzy Simulation Approach, I*nternational Conference on Computer Modeling and Simulation (ICCMS '09)*. pp: 88 - 92.

Kumar, R. Grover, P.S. Kumar, A., 2010. A Fuzzy Logic Approach to Measure Complexity of Generic Aspect-Oriented Systems, *ETH Zurich,* **9**, (3), pp. 43 - 57.

Chang, E. Dillon, TS. Cook, D., 1997. An Intelligent System Approach to Usability Evaluation, *IEEE Conference on Intelligent Information Systems*, pp. 218 - 226.