USELEARN: ADAPTIVE USABILITY EVALUATION OF E-LEARNING SYSTEMS

by

Asil ÖZTEKİN

A thesis submitted to

the Graduate Institute of Sciences and Engineering

of

Fatih University

in partial fulfillment of the requirements for the degree of

Master of Science

in

Industrial Engineering

June 2006
Istanbul, Turkey

APPROVAL PAGE

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

____
Prof. Dr. Mazhar ÜNSAL
Head of Department

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

____
Prof. Dr. Alexandar NIKOV
Supervisor

Examining Committee Members

Prof. Dr. Alexandar NIKOV ______________________

Assoc. Prof. Selim ZAİM ______________________

Assist. Prof. Nurullah ARSLAN ______________________

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

____
Assist. Prof. Nurullah ARSLAN
Director

Date
June 2006

USELEARN: ADAPTIVE USABILITY EVALUATION OF E-LEARNING SYSTEMS

Asil ÖZTEKİN

M. S. Thesis - Industrial Engineering
June 2006

Supervisor: Prof. Dr. Alexandar NIKOV

ABSTRACT

UseLearn approach for adaptive usability evaluation of eLearning systems is developed. Based on the neural networks model, usability checklist dimensions are adaptively selected and relevant data are collected for usability evaluation of an e-learning system. A case study for usability evaluation of eLearning system in cell biology by *UseLearn* in Fatih University and Fatih College is carried out. The analysis of experimental results showed that *UseLearn* approach supports allocation of usability problems and defining relevant improvement measures. The main advantage of *UseLearn* is the adaptive selection of most significant checklist dimensions and thus reducing the time for usability evaluation.

Keywords: eLearning, Usability, Principal Components Analysis, Neural Networks, Pareto Analysis.

E-ÖĞRENME SİSTEMLERİNİN ADAPTİF KULLANIŞLILIK DEĞERLENDİRMESİ

Asil ÖZTEKİN

Yüksek Lisans Tezi – Endüstri Mühendisliği
Haziran 2006

Tez Yöneticisi: Prof. Dr. Alexander NIKOV

ÖZ

E-öğrenme sistemlerinin adaptif kullanışlılık değerlendirmesi için *UseLearn* yaklaşımı geliştirilmiştir. Sinir ağları modeline bağlı olarak kullanışlılık anket maddeleri adaptif olarak seçilmiş ve bir e-öğrenme sisteminin kullanışlılık değerlendirmesi için gerekli veri toplanmıştır. *UseLearn* metodu vasıtasıyla Fatih Üniversitesi ve Fatih Koleji'nde hücre biolojisi ile ilgili e-öğrenme sisteminin kullanışlılık değerlendirmesi için bir vaka analizi yapılmıştır. Deneysel sonuçların analizi *UseLearn* yaklaşımının kullanışlılık problemlerini belirlemeyi ve ilgili geliştirme ölçütlerini tanımlamayı desteklediğini göstermiştir. *UseLearn* yaklaşımının ana avantajı en önemli anket maddelerinin adaptif seçimi ve böylece kullanışlılık değerlendirmesi için harcanan zamanı azaltmasıdır.

Anahtar Kelimeler: e-öğrenme, Kullanışlılık, Ana Bileşenler Analizi, Sinir Ağları, Pareto Analizi.

To my family and friends,

ACKNOWLEDGEMENT

I would like to thank my thesis supervisor Prof. Dr. Alexander NIKOV for his guidance throughout the research, for creating new ideas to contribute to this thesis and his comments on analyzing the experimental case study.

I would like to thank also to biology teacher Dr. İrem İffet Uzonur for her help collecting data in Fatih College and research assistant Mustafa Ulaşlı for his help collecting data in Fatih University, Department of Biology. I present my sincere appreciation to the Assist. Prof. Ferhat Karaca and Lecturer Ali Türkyılmaz for their invaluable help during the implementation phase of the thesis.

Thanks go to the jury members Assoc. Prof. Selim Zaim and Assist. Prof. Nurullah Arslan for their valuable suggestions and comments during the graduate seminar standing.

TABLE OF CONTENTS

LIST OF TABLES

TABLE

LIST OF FIGURES

FIGURE

CHAPTER 1

INTRODUCTION

Usability evaluation is an increasingly hot topic of human factors engineering/ergonomics. As the web based systems are used more broadly in every area nowadays, similarly learning is also performed more often by the help of the internet, which is called as eLearning.

As mentioned by Nielsen (1994), usability is related to quality. Which is including the other one is a debate that can change from one study of area to the other one. It doesn't matter either usability is the subset of quality or the opposite or they have some common overlapping items and each has differentiating items at the same time. In this thesis in Chapter 2 we search for the answer of this question for eLearning systems evaluation. We compare the usability and quality valuation checklists; select the most comprehensive and the best fitting one to our case study. These quality and usability checklists are compared and the overlapping and differentiating items are found. By this way, a modified checklist of simultaneous usability and quality checklist is developed.

In Chapter 3, an adaptive usability evaluation method which we named as UseLearn algoritm is developed. UseLearn is a well-established method including principal components analysis, neural networks and Pareto analysis. In this method it is proposed that, instead of dealing with all usability problems to get a more usable eLearning system, it is more logical to follow a quantitative method by using principal component analysis to aggregate the usability checklist questions. After that, we can apply neural networks to find out the most critical dimensions causing main usability problems. By neural networks, the weights of each dimension are determined. But here comes the question "Which dimensions are the most critical to improve the usability?" At this moment, UseLearn algorithm suggests to apply

Pareto analysis to select the most critical dimensions affecting usability problems significantly. It is critical here to take care what is adaptive. In *UseLearn*, usability evaluation is adaptive depending on the analysis and the commends not the eLearning system.

In Chapter 4, we performed a case study with the data gathered in Fatih University and Fatih College by using the eLearning system *moodle* in a cell biology course. Using this data, UseLearn algorithm is applied and usability problems are revealed. Depending on these usabilty problems and the comments made by the test participants, improvements are done and the eLearning system is improved and the participants are retested. Just focusing on a few usability problems it is seen that usability satisfaction of the participants are increased as it is proposed in our UseLearn model.

CHAPTER 2

CRITICAL OVERVIEW OF USABILITY EVALUATION OF E-LEARNING SYSTEMS

2.1 E-LEARNING SYSTEMS

2.1.1 Definition of eLearning

eLearning stands for Electronic Learning. It is a formalized teaching and learning system specifically designed to be carried out remotely by using electronic communication. This approach is nowadays realized all over the world. The fields of eLearning widely vary.

They include:

- Students teaching
- Integration of learning programs to LMS
- Knowledge control
- Staff training
- Products presentation

Popular eLearning technologies are:

- Voice-centered technology, such as CD or MP3 recordings or Web casts
- Video technology, such as instructional videos, DVDs, and interactive videoconferencing
- Computer-centered technology delivered over the Internet or corporate Intranet

Studies indicate that distance learning can be as effective as the traditional format when the methods are appropriate to the teaching tasks, there is student- teacher interaction, and the teachers provide students with appropriate and timely feedback.

For purposes of this case study, the term eLearning is used to mean learning using both a computer and the Internet. The term eLearning is, in some senses, a synonym for Open and Distance Learning (ODL). eLearning products or services take various forms. They may be single courses and/or entire programs; entire courses and/or course units, lessons or components; or elements of an eLearning package, e.g. a learning management system or an ePortfolio tool. The eLearning may be:

- offered for credit at an education institution and/or for general interest without credit;
- aimed at individuals or entire groups in classes;
- targeted to specific age groups and/or any age group;
- offered by public and/or commercial education and training agencies.
- From the purchaser perspective, the eLearning service may:
- provide instruction with or without various support services, such as access to a library
- be very expensive or free of charge;
- be really effective or of questionable quality.

Both providers and consumers of eLearning want education and training products and services that are effective and efficient. The term quality is used to encompass these concepts. Consumers of eLearning may be individual students, schools boards, education and training departments of governments, corporations. Providers may be publicly-funded schools, universities and colleges, or they may be private enterprises producing portions of eLearning content, design and production, delivery and management of learning, and/or student management (Barker,2004).

2.1.2 eLearning Styles

There are three main types of eLearning systems include:

1. Self-Directed Learning:

Student acts alone and works through the materials that are delivered to him/her over the internet.

- There is no instructor or group of peer students to communicate with.

2. Synchronous Learning:

Instructor and all the students meet over the internet at the same time and are connected by technology that acts like real-time interactive audio and video.

- All on-line participants can interact together

3. Asynchronous Learning:

Instructors and students do not meet at exactly the same time. Instructor and students interact by exchanging messages (bulletin board or team room)

Asynchronous systems, which are heavily reliant on print and text, are good for conveying large quantities of information, usually via either electronic or land mail systems (Bates,1995). Because asynchronous systems slow down the two-way communication process, they offer learners the flexibility and convenience of participating in learning activities at times and places best suited to their needs (Daniel, 1996). Web-based courses are an effective application of asynchronous distance education; they provide access to a broad range of learning resources and learner assistance since on-line help and tutorial support can be built into the programs (Smith and Dillon,1999). Furthermore, asynchronous Web courses enable learners to reflect on and interact with new information before having to respond to it. An additional benefit is that such systems allow the instructor to give each student a high degree of individualized attention (Moore and Kearsley,1996).

Synchronous learning systems, such as audio-conferencing and videoconferencing, have different strengths. The highly structured learning environment and the teaching-by-telling method, reminiscent of traditional education and training, make interactive synchronous learning especially useful for reinforcing knowledge or for speedy correction of misunderstandings. Group-based synchronous classes enable instructors to pace learning activities consistently, and they may improve motivation in some learners through the use of already-familiar teaching methods. Synchronous communication is particularly appropriate when learners and instructors in different locations need to work collaboratively or to conduct discussions within a fixed timeframe (Whitworth, 1999). Supporting these activities offers a powerful benefit since they enable learners to be part of a social and socializing environment (Hiltz,1995).

Other distinguishing features of asynchronous or synchronous systems influence their relative effectiveness and therefore need to be actively planned for and/or managed. Asynchronous systems, for example, depend on students' self-discipline in participating and maintaining an appropriate pace. Additionally, the 24/7 availability of communication in asynchronous systems can tempt instructors to present excessive amounts of information and tasks. Similarly, students can easily swamp instructors with masses of e-mail that seem to demand an immediate response. In synchronous learning environments, on the other hand, the transience of the activities requires that learners be able to take good notes and recall salient points from the discussions. Unlike learners in asynchronous systems, which are usually supported by comprehensive educational materials available for repeated use, those in synchronous systems must process and understand the presentations and discussions as they happen (Bates, 1995).

Some media are able to support both asynchronous and synchronous distance education. Video- and audio-based media can be used asynchronously (e. g., video- and audio-taped instruction) or may be combined with telephone and computers to enable real-time conferencing. Because large groups of students are likely to exhibit wide variability in their characteristics, the use of a combination of media for instruction increases the chances of positive learning outcomes by increasing the range of learning styles that can be accommodated (Moore and Kearsley, 1996).

Although there is a tendency to think that "high tech" solutions will offer high levels of effectiveness, evidence from research suggests that simpler technologies, when combined with well-designed instruction, are as—and sometimes more—effective than more complex and expensive technologies (McIsaac and Gunawardena, 1996). In other words, factors other than the complexity of the delivery system are the real determinants of effectiveness. Student characteristics, program design, content, instruction, and program administration all have a major impact on effectiveness.

Developing a clear conception of the goals and objectives of a program prior to the selection of a particular learning environment and/or supporting technologies is the first step in maximizing effectiveness (Bates, 1995 - Dede, 1996). These goals then can be used in establishing the appropriate balance of elements that, for any given learning situation, will result in positive learning outcomes and overall effectiveness.

2.1.3 Comparison of Traditional Learning and eLearning

Reading Skills: browsing/scanning skills (particularly important when using the World Wide Web to locate relevant websites)

Writing: Summarizing the key points referencing information keeping records

Research Skills: searching skills are essentially a part of research skills

The table is not intended to cover all the sub-skills, but merely to provide an initial basis for comparison.

The two major differences between traditional and eLearning skills are the context and degree of importance of the skill. eLearning may be learning as a distance for both their peers and tutors, so they need to be far more self-sustained than the traditional learner. Traditional learning provides many informal opportunities, such as a brief chat in the corridor with other learners, to discover that views, whereas eLearning requires you to send an email, a more formal activity. Learners have had years of practice in face-to-face communicating but most will be relatively inexperienced at being dependent on short written messages

In face-to-face communications you can see the facial expressions of people, hear the tone of voice and listen to the words used. In eLearning you only have the written words to communicate through and this notoriously leads to misunderstandings, It is more difficult to convey precise meaning. Various ways have been developed to help convey emotions such as the use of emotions, a code based on punctuation or other symbols such as:

- ☺ happy
- ☹ sad
- Using capitals or upper case means you are shouting.

However only a minority of email users includes emotions and they are inevitably basic in comparison to the non-verbal communication that everyone has grown up using. They may be confusing if the person receiving the message does not understand their purpose.

Traditional learners have the benefit of a tutor's judgment in observing their group and realizing that some individuals are confused by the topic or that others need to be reminded about a forthcoming test. An eLearning tutor has far less information on which to base a judgment, so the learners need to be more self-reliant. Time management skills become more important since it is needed to be in control on learning. Atutor or peers can not be relied on informally reminding the deadlines. In later chapters, the opportunity to develop these skills will be given. (Clarke, 2004)

eLearning assumes the user is a component and confident user of computers and communication technology. In table 2.1, Clarke summarizes the comparison of traditional learning and eLearning.

Table 2.1 Comparison of Traditional Learning and eLearning (adapted from Clarke, 2004)

Traditional skills	**eLearning skills**	**Difference**
Time management	Time management	Time management is critical in eLearning since it provides greater opportunity to take control of your own learning. This is also true of other forms of open and distance learning
Acceptance of responsibility	Acceptance of responsibility	eLearning provides more opportunities for learners to take responsibility for their learning than traditional learning does. This is also true of other forms of open and distance learning
Planning	Planning	The benefits of eLearning include giving learners more freedom to choose when and how they study, so placing on them emphasis for planning. Traditional courses are often determined by the tutor and accompanied by timetables and study guides. This is also true of other forms of other forms open and distance learning.
Searching skills libraries	Searching skills world wide web	Scale –the world wide web is enormous in comparison to any physical library
Assessing quality- written and other	Assessing quality world wide web	The world wide web has few quality insurance mechanisms. Books and

physical content		other printed educational content have established means of judging quality. Anyone can launch a website, but producing a textbook requires the agreement of publishers, peers and reviewers.
Listening- to peers and teachers during presentations and discussion	Listening is required only occasionally e.g. when the program is based on video	Listening is a key skill in most forms of traditional learning, while it is frequently only plays a minority role or none at all in eLearning
Reading mainly printed material	Reading is a key skill in eLearning. Most information is presented as text displayed on a screen	Since a majority of the material is text, reading is a key skill in eLearning comparable to the roles reading and listening play in traditional learning Browsing is the normal way that the content of websites is read to locate relevant content
Writing-mostly in the form of note taking or completing exercises	Writing (keyboard skills) for communications, note taking and exercises	Writing (keyboard skills) is essential for eLearning communication (e.g. email) as well as note taking and exercises. Writing is the main online communication method
Self-assessment	Self-assessment	This is a key skill in all forms of learning. In traditional learning there are many opportunities to compare your performance with your peers.
Collaborating with others face to face	Collaborating with others through communication software (e.g. email)	The key difference is time. A face to face group will often agree regular meetings so that tasks are achieved quickly. Online group members will each have their own time scales and may well live in different time zones, so that collaboration often spread over a long period. Motivations is sometimes difficult to maintain
Problem solving individually or small groups	Problem solving individually or with a group at a distance	The significant difference when working with others is that in eLearning they are at a distance and it is therefore difficult to judge their views.

2.1.4 Benefits of eLearning

Some advantages of eLearning can be summarized like the following items:

- It is less expensive to support
- It is less expensive to produce
- It is not constrained by geographic considerations
- It is more flexible in terms of time
- It can be delivered virtually anywhere
- The number of students is never limited by the size of the classroom
- It can work from any location and any time
- The teacher is able to monitor student progress
- The individual learning can be easily realized

Main benefits as an e-learner are that you have considerable freedom of choice over:

- Place
- Pace
- Time

(Kruse, 1999) summarizes the benefits of eLearning as shown in table 2.2.

Table 2.2 Benefits of eLearning (adapted from Kruse, 1999)

Benefits of eLearning	
Benefits	**Description**
Technology has revolutionized business; now it must revolutionize learning.	The need to transform how organizations learn points to a more modern, efficient, and flexible alternative: eLearning. The mission of corporate eLearning is to supply the workforce with an up-to-date and cost-effective program that yields motivated, skilled, and loyal knowledge workers.
Anywhere, anytime, anyone.	We estimate that approximately 80% of the professional workforce already uses computers on the job. Technical obstacles, such as access, standards, infrastructure, and bandwidth, will not be an issue two years from now. The growth of the World Wide Web, high-capacity corporate networks, and high-speed desktop computers will make learning available to people 24 hours a day, seven days a week around the globe. This will enable businesses to distribute training and critical information to multiple locations easily and conveniently. Employees can then access training when it is convenient for them, at home or in the office.
Substantial cost savings due to elimination of travel expenses.	The biggest benefit of eLearning, however, is that it eliminates the expense and inconvenience of getting the instructor and students in the same place. According to Training Magazine, corporations save between 50–70% when replacing instructor-led training with electronic content delivery. Opting for e-training also means that courses can be pared into shorter sessions and spread out over several days or weeks so that the business would not lose an employee for entire days at a time.
Just-in-time access to timely information.	Web-based products allow instructors to update lessons and materials across the entire network instantly. This keeps content fresh and consistent and gives students immediate access to the most current data. Information can be retrieved just before it is required, rather than being learned once in a classroom and subsequently forgotten. Training Magazine reported that technology-based training has proven to have a 50–60% better consistency of learning than traditional classroom learning (c-learning).
Higher retention of content through personalized learning.	Since they can customize the learning material to their own needs, students have more control over their learning process and can better understand the material, leading to a 60% faster learning curve, compared to instructor-led training. The delivery of content in smaller units, called "chunks," contributes further to a more lasting learning effect. Whereas the average content retention rate for an instructor-

	led class is only 58%, the more intensive eLearning experience enhances the retention rate by 25 – 60%.
Improved collaboration and interactivity among students.	Teaching and communication techniques which create an interactive online environment include case studies, story-telling, demonstrations, role-playing, simulations, streamed videos, online references, personalized coaching and mentoring, discussion groups, project teams, chat rooms, e-mail, bulletin boards, tips, tutorials, FAQs, and wizards. Distance education can be more stimulating and encourage more critical reasoning than a traditional large instructor-led class because it allows the kind of interaction that takes place most fully in small group settings. Another study found that online students had more peer contact with others in the class, enjoyed it more, spent more time on class work, understood the material better, and performed, on average, 20% better than students who were taught in the traditional classroom.
Online training is less intimidating than instructor-led courses.	Students taking an online course enter a risk-free environment in which they can try new things and make mistakes without exposing themselves. This characteristic is particularly valuable when trying to learn soft skills, such as leadership and decision-making. A good learning program shows the consequences of students' actions and where/why they went wrong. After a failure, students can go back and try again. This type of learning experience eliminates the embarrassment of failure in front of a group.

The user is potentially free to study at any location he/she wants. So if he/she likes to work at home, he/she can combine studying with family responsibilities or avoid the frustrations of commuting.

The tutor and other learners in the class or group often set the pace of studying. eLearning gives the user to choice of how quickly or slowly to learn. If he/she wants to work through the night, he/she can or if he/she likes to fit short bursts of activity into his timetable to allow for taking care of elderly parents or children, he is free to do so.

All traditional courses have a fixed timetable of classes and activities around which he/she must work. eLearning provides him with a considerable degree of choice. He can study in the middle of the night or during the day, whatever is best for him.

2.2 E-LEARNING QUALITY EVALUATION APPROACHES

The role of quality is to ensure that interdependent processes are properly coordinated toward predetermined goals. Quality systems must be created in the context of a mission statement, set of objectives, and core values. The needs of stakeholders who are affected by the activities of the system also need to be considered.

Quality assurance is a journey rather than a destination. By providing clear quality assurance procedures that considered factors such as institutional strategy and context, the levels of eLearning and the imperative of flexible usefulness, this is in an excellent position to consistently develop high-quality eLearning solutions and learn from its experience.

There are many ways in which eLearning can be applied in tertiary education.

Because the applications can be technologically complicated, and because their use does not always match well with traditional modes of teaching and learning, much care needs to be taken in the design, creation and implementation of eLearning solutions (Nichols,2002).

While quality is difficult to define, its importance is universally appreciated (Garvin, 1988). Quality's commercial importance comes from its perceived ability to lower costs, improve employee commitment, and ensure continuous improvement within a dynamic environment (Dawson and Palmer, 1995). Quality is not just about zero defects; improving the performance and style of an end product are also important factors (Deming, 1994). Garvin (1988, p.36) states the role of quality personnel:

Today's quality professionals bear little resemblance to their turn of the century predecessors.They are managers, not inspectors; planners, not controllers; sensitive to markets as well as to manufacturing.

Quality is described as a concept rather than a technique, so its implementation is very much dependent on the type of organisation or process at hand (Gilmour and Hunt, 1995). Identifying processes is an important step toward improving them and predicting the consequences of changes; process maps should consider all aspects of the service including suppliers, clients, design, production, and delivery (Deming, 1994; Gilmour and Hunt, 1995). According to Deming (1994), some 94% of quality problems result from a faulty system.

A quality system must be based on an understanding of interdependence, that is, the reliance of the overall process on the effective performance of each and every task. Deming defines a system as "a network of interdependent components that work together to try to accomplish the aim of the system" (1994, p.50). It is vital that every task in a process be performed properly and that all relationships between different tasks are understood. The quality of the finished product is the direct result of the quality throughout the process used to create it. A problem at any stage in the process will affect the quality of the entire process. Effective management of an integrated process is key to ensuring quality outcomes. A quality system does not just aim to meet the needs of clients; other stakeholders such as employees and shareholders must also be considered, and the system must be compatible with the organisation's overall strategic direction (Gilmour and Hunt, 1995). Flow charts assist in the understanding of a system (Gilmour and Hunt, 1995).

There are many different approaches to quality (Walklin, 1992; Hagar, 1998), most of which are applied at the organisational level rather than that of individual modules or projects. Quality assurance is one approach, defined by Gilbert (1992, p.32) as "the assembly of all functions and activities that bear upon the quality of a product or service so that all are treated equally, planned, controlled and implemented in a systematic manner." More specifically, a Quality Assurance (QA) system documents procedures with the aim of ensuring that the overall process meets specified objectives and to demonstrate that quality is a managed outcome (Dawson and Palmer, 1995). As such it is a sub-function of Total Quality Management (TQM), which is more institutional in its application. QA is an activity; TQM is a philosophy.

Walklin (1992) suggests the establishment of a mission statement and a set of objectives as the basis for a customer focus that will in turn provide the basis for a quality programme. Generating a quality policy assists in the adoption of a TQM mindset within an organisation (Gilbert, 1992). Further, quality systems should aim at continuous self-improvement (Walklin, 1992).

Quality is made up of many elements. For eLearning products, the following is suggested as quality criteria (based on Garvin, 1988):

• **Performance –** the finished product should operate in an effective way, as determined by the end-user.

• **Features –** the 'bells and whistles' incorporated into the finished product should be appropriate, and not detract from the overall objectives of the project.

• **Reliability –** the finished product should not be subject to malfunction.

• **Conformance –** the finished product should comply with industry standards, using standard technologies (though those technologies can be pushed to their utmost) and reflect establishededucation theory.

• **Durability –** the finished product should be relevant and either timeless (in the case of teaching established principles) or easily updated.

• **Serviceability –** it should be easy to repair or adjust the finished product as required.

• **Aesthetics –** the overall 'feel' of the finished product should be professional and user-friendly.

• **Perceived Quality –** the finished product should enhance the reputation of UCOL as a quality eLearning provider.

Development of the eLearning QA system required firstly a firm set of responsibilities and activities performed by the eCampus team. Once these were identified, quality assurance processes were created to make sure that various quality outcomes were met during development. Foundational to these processes is a set of aims, objectives and core values.

In another approach as mentioned in (Wilkinson, 2005), a holistic approach for eLearning quality is developed. This approach is proposed as shown in the figure 2.1.

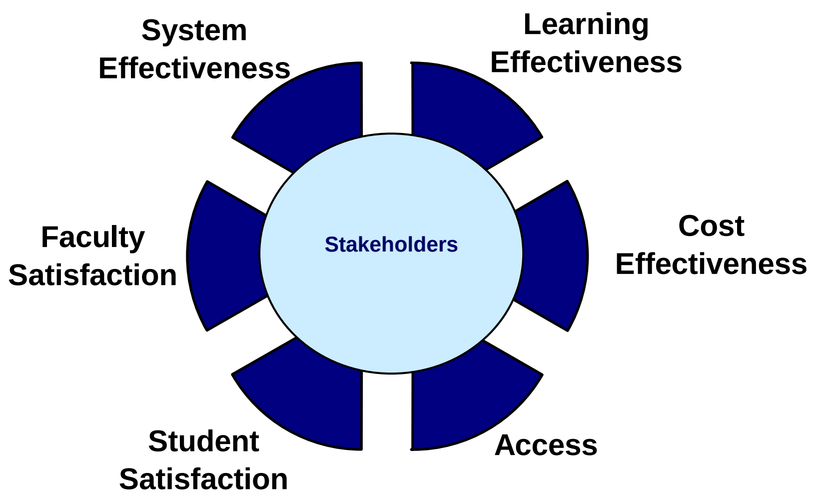

Figure 2.1 Holistic Approach to eLearning Quality (adapted from Wilkinson, 2005)

Moore (2004) mentioned the Sloan-C and explains this graphic as the following and makes a formulation for its explanation. It keeps in mind that quality is a work in progress and each organization seeks to measure quality in terms of its own distinctive, dynamic mission and the people who embody it. Thus, the Sloan-C quality framework enables each organization to set its own standard for each pillar. For example, a school could weight the importance of each measure in the following equation:

Quality = k1 *Learning Effectiveness + k2 *Cost Effectiveness and institutional commitment +k3 *Student Satisfaction + k4 *Faculty Satisfaction + k5 *Access

For a selective admissions school, k5 *Access might not be as important as it is for open admissions schools. Clearly, an organization can take different looks at the scales (for example an organization could assess its Learning Effectiveness on a scale comparing it to others in the National Study of Student Engagement, or Cost Effectiveness compared to US rankings in news reports, or even Student Satisfaction according to MSN's Best Party Schools).

Measures of quality begin with vision and mission. For each of the pillars, the statements below describe an ideal environment:

Learning effectiveness

(Moore, 2004) demonstrates that the quality of learning online is comparable to the quality of its traditional programs:

- Interaction is key: with instructors, classmates, the interface, and via vicarious
- Online course design takes advantage of capabilities of the medium to improve learning (testing, discussion, materials)
- Communications and community building are emphasized
- Swift trust characterizes the online learning community
- Distinctive characteristics of programs are highlighted to demonstrate improved learning
- On-campus and online instruction achieve comparable learning outcomes, and the institution ensures the quality of learning in both modes with metrics tracking instructional methods, student constituencies and class size

Cost effectiveness and institutional commitment

Institutions continuously improve services while reducing cost

- Cost effectiveness models are tuned to institutional goals
- Tuition and fees reflect cost of services delivery
- Scalability, if an institutional objective, can be accommodated.
- Partnering and resource sharing are institutional strategies for reducing costs
- Mission-based strategies for cost reduction are continuously formulated and tested
- Intellectual property policies encourage cost effective strategies
- Traditional programs

Access

All learners who wish to learn online have the opportunity and can achieve success

- Diverse learning abilities are provided for (at-risk, disabilities, expert learners)
- The reliability and functionality of delivery mechanisms are continuously evaluated
- Learner-centered courseware is provided
- Feedback from learners is taken seriously and used for continuous improvement

- Courses that students want are available when they want them
- Connectivity to multiple opportunities for learning and service is provided

Faculty Satisfaction

Faculty achieve success with teaching online, citing appreciation and happiness

- Faculty satisfaction metrics show improvement over time
- Faculty contribute to, and benefit from online teaching
- Faculty are rewarded for teaching online and for conducting research about improving teaching online
- Sharing of faculty experiences, practices and knowledge about online learning is part of the institutional knowledge sharing structure
- There is a parity in workload between classroom and online teaching
- Significant technical support and training are provided by the institution

Student Satisfaction

Students are successful in learning online and are typically pleased with their experiences.

- Discussion and interaction with instructors and peers is satisfactory
- Actual learning experiences match expectations
- Satisfaction with services (advising, registration, access to materials) is at least as good as on the traditional campus
- Orientation for how to learn online is satisfactory
- Outcomes are useful for career, professional and academic development (Moore, 2004)

In (Stephenson, 2005) he mentioned different levels of eLearning quality as briefly explained below:

The broader perspective

Lopez et al at an ESOE forum in 2003 suggested gathering data in seven areas reflecting a Total Quality Management (TQM) perspective on ICT based programmes. Their approach embraces institutional characteristics and programme structures as well as learning activities. Though not explicit, Lopez's indicators imply a learner managed learning pedagogical approach. The seven dimensions are:

- availability of ICT qualifications
- online management of academic information
- on-line self-inscription
- personalisation of online facilities
- virtual university campuses
- proportion of eLearning experiences
- participation in shared virtual campus experiences.

At the operational level

Indiana University's Center for Research on Learning and Technology (CRLT) also has a list of seven criteria but in contrast to Lopez et al's TQM approach, theirs is more explicitly related to identifying good practice within a conventional pedagogical model that emphasizes good practice within a fairly traditional teaching paradigm. Good ICT practice, according to (Grahan et al, 2002):

- encourages student-faculty contact
- encourages cooperation among students
- encourages active learning
- gives prompt feedback
- emphasizes time on task
- communicates high expectations
- respects diverse talents and ways of Learning

At the pedagogical level

An early example focusing on pedagogical aspects of quality is (Reeves, 1997) list of fourteen pedagogical dimensions that should be taken into account when evaluating Computer Based Education (CBE). Reeve's fourteen dimensions were:

- epistemology

- pedagogical philosophy
- underlying psychology
- goal orientation
- experiential value
- teacher role
- program flexibility
- value of errors
- motivation
- accommodation of individual differences
- learner control
- user activity
- cooperative learning
- cultural sensitivity

The underlying pedagogical stance of Reeve's list is constructivist and learner centred. Quality, Reeves argues, needs to be discussed in the context of the preferred pedagogical paradigm. CBE, Reeves sadly observed, is more frequently seen as 'an alternative delivery system for traditional pedagogy rather than as a tool for implementing alternative pedagogical dimensions. Evaluation approaches based upon clearer delineation of the pedagogical dimensions within different types of CBE' he further argued, 'will surely be a step forward.

In a workshop called eLearning Quality Euler&Seufert mention the added-value of eLearning environment as the didactic potential of eLearning in different aspects of view and summarize them as shown below.

Aspects:

- clear, comprehensive illustration of learning content
- anonymity: incentive for exploration
- individualization of learning:
- independent of time and space
- selection of learning content
- duration and speed of learning
- asking for tutorial support

- more time for challenging learning goals in face-to-face-sessions
- new options for interaction among learners
- IT both method and content

While they are mentioning the areas for implementing a sustainable eLearning strategy, they also figurize them as shown in the figure 2.2.

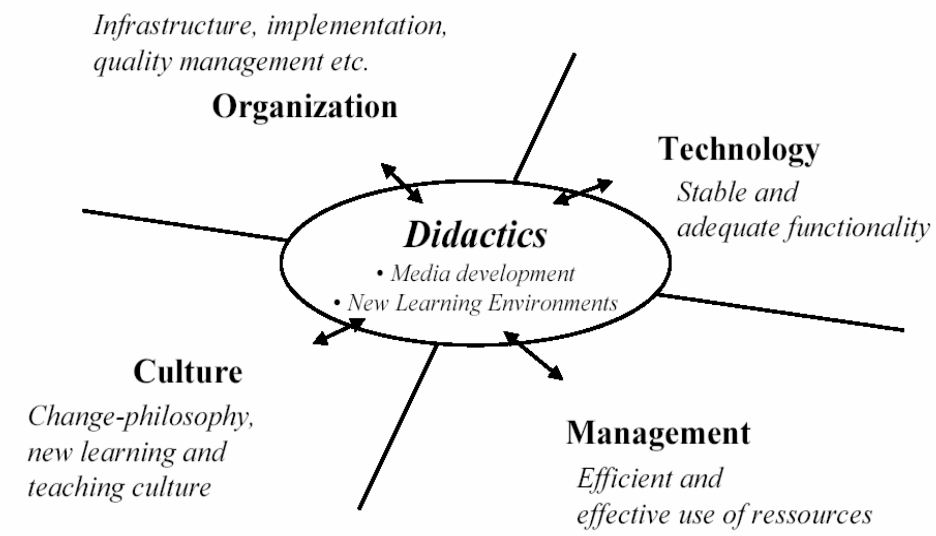

Figure 2.2 Areas for Implementing a Sustainable eLearning Strategy (adapted from Euler and Seufert, 2005)

Pawlowski and Ehlers mention some quality standards in CEN/ISSS Workshop Learning Technologies in 2005 in Helsinki like those and claim that promoting eLearning quality standardization is possible by promoting:

- diversity&flexibility
- solutions to improve transparency
- adaptations of international standards
- contributing to international standards
- improving a consensus

In Learning Innovations Forum (Barker, 2004) mentioned that the Quality Standards are consumer-oriented, comprehensive, consensus-based, adaptable, futuristic and recommended only. The entire set is lengthy and complex.

In brief, the standards begin with what is most important to consumers – assurance that they will learn content skills and knowledge that are relevant and recognized together with lifelong learning skills that are transferable and applicable. When consumers are assured their investment of time and finances will be rewarded with recognized competencies and credits, they then concern themselves with the details of student services and delivery – the teaching, learning, assessment and support processes and practices. When they are assured that teaching and learning are appropriate and effective, they finally concern themselves with the nature of the organization standing behind the learning service – the quality of staff, budgets and plans. In summary, the Open eLearning Quality Standards are comprehensive of all elements of the system: outcomes and outputs, processes and practices, inputs and resources.

Faced with the variability of cost, quality and innovation in eLearning, FuturEd Inc.4 hypothesized that consumer-based quality guidelines for learning technologies could be used by all stakeholders in eLearning.

- Students and purchasers could use them to make informed choices.
- Producers of eLearning could use them to develop, evaluate, improve and market their products and services.
- Policy makers would use them to understand the needs and perspectives of the public, and the fit between what is needed and what is available.
- The ODL community could use the quality standards to ensure quality in transnational learning products and services.
- The international marketplace could use Canadian standards to appreciate Canadian quality values in technology-assisted learning products and services.(Barker,2004)

When a new technology enters the - already crowded - communication market, the question on its effectiveness and efficiency usually arise immediately, dividing its (possible) target public into different parties, according to different judgements. It is not, of course, the first time that education is challenged by technologies: on the contrary, it is usually a major test-bed for new “technologies of the word” (Bolter, 2001): here, people look for a confirmation of the social relevance of their innovations, as well as for economic investments endorsed by the social community.

While, at first, the attention of researchers was attracted by the issue of effectiveness of the use of new ICT, by comparing “traditional” courses with courses using those technologies

(Phipps & Merisotis,1999), more recently it is becoming clear that the relevant question is not mainly if eLearning is effective or not, but whether it is efficient, and under what conditions.

A wide corpus of research projects has shown quite clearly that eLearning can be as effective as Learning activities without the "e" prefix, this kind of research has thus met that big impasse called "the not significance phenomenon" by Russell (1999).

A more promising research area seems to be that of teachinglleanning quality: the problem is not any more whether or not to use eLearning, but how to implement it to offer a high quality learning experience (Phipps & Merisotis 2000; Peters 2002). Universities are among the institutions more studied and researched (Rowley *et al.* 1998), due to the fact that they implemented eLearning quite early, are relatively easy to study and willing to be studied, and because it is quite easy to gather information (a different situation is found in for-profit organizations).

Awareness of the quality issue is growing at the same time as the awareness of the management issue: the more eLearning enters an organization - leaving the status of trial or "technological gadget" and becoming an integral part of its educational activities - the greater the importance of its harmonization and deep integration with the institutional structure, its procedures and practices, its ***myths*** and ***rites.*** Again, the ecological perspective taken by diffusion theories can help better understand what is happening where ekarning starts to become flesh of an educational organisation: its nature changes somehow, everything has to be accommodated accordingly and nothing remains exactly the same.

Their research goes in this direction, trying to assess how some European Universities have integrated eLearning activities into their educational offer, and how do they perceive the quality of their own eLearning initiatives.

Georgina Stein handles with eLearning Quality in her study called "Putting the e into Learning&Teaching: Quality Assurance Guidelines". She categorizes and makes groups as follows:

Quality & Effectiveness

- **The eLearning Model –** eLearning Models vary greatly depending on the nature of the course and the delivery organisation; eLearning might include face-to-face delivery plus the use of email to communicate with learners, posting notes to a web page or by using a course management system which includes asynchronous and synchronous discussion boards as well as other course management and communication tools.

- **The eTeaching Strategy -** eTeaching involves using online tools and resources to support learning as described above.

Quality & Support

- **QA Strategies and Procedures**
- **Operational Management and Administration**

In details these items and their subdimensions are expressed as shown below and the checklist questions are stated.

The eLearning Model – *Technical Factors*

- **Organisation Factors –** Technical**:** How do you ensure that the eLearning course is in an appropriate format and is accessible to all learners?

Connectivity: Is there sufficient bandwidth to implement and sustain eLearning?

Hardware: Are there sufficient accessible computers available to implement and sustain eLearning?

*Software***:** Are the relevant eLearning and eTeaching technologies available to meet the course requirements?

*Resource Location***:** How do you ensure that the eLearners and eTeachers are able to access appropriate resources?

*Course Evaluation & Feedback***:** What strategies are in place for evaluating the technical infrastructure and support systems?

The eLearning Model – *Academic Factors*

- **Organisation Factors –** Academic**:**

*Course Design & Structure***:** Is the eLearning course designed and structured to meet the aims, objects and the general requirements of the course?

*Course Content***:** Is the eLearning content in an appropriate format? Have the eLearning resource materials been developed to meet a specific quality assured set of standards?

*Course Delivery***:** How does eLearning enhance course delivery?

*Course Evaluation & Feedback***:** What systems are in place for evaluating the quality of the academic input into the course and how is this fed back to the organisation centrally?

The eLearning Model – *eLearner Factors*

- **eLearner Factors:**

*Recruitment Strategy***:** How do eLearners know they are ready to embark on an eLearning course?

*Entry Characteristics***:** What entry characteristics do eLearners need to have to enable them to successfully participate in the eLearning course? E.g. Personal skills and competence.

*ICT Skills***:** Do all eLearners have the necessary ICT Skills to access eLearning? What support is available to them?

*Roles & Responsibilities***:** Do all eLearners know and understand what they are required to do to gain maximum benefit from the course to meet the course eLearning requirements?

*Course Evaluation & Feedback***:** Do all eLearners know what systems in place for evaluating the course [technical and academic] and how this is fed back to the organisation?

The eTeaching Strategy – eTeaching Factors

- **eTeaching Factors:** How do you ensure that all eTeaching meets a set of quality assured standards?

*Teacher Recruitment to eLearning***:** What methods do you use to recruit or induct eTeachers in order to ensure that they have suitable eLearning experience and expertise?

*ICT Skills***:** Do all eTeachers have the necessary ICT Skills to deliver and support eLearning? What support is available to them?

*Roles & Responsibilities***:** How do you ensure that all those involved in the development and delivery of the eLearning course understand their roles and responsibilities?

*eTeaching Training***:** What teacher eTraining programmes are available?

*eTeaching Performance***:** How do you evaluate the performance of individual eTeachers?

*Continuing Professional Development***:** How do you identify future eLearning training and development needs in order to ensure that all eTeachers have relevant up-to-date knowledge, skills and understanding?

*Course Evaluation & Feedback***:** Do eTeachers know what systems are in place for evaluating the course and how this is fed back to the organisation centrally?

QA Strategies & Procedures

- **Quality Assurance Factors:** How do you ensure that all parties involved in the eLearning course have a shared understanding of quality assurance and quality improvement structures and procedures?

*Assessment & Recording***:** To what extent are all eTeachers and eLearners involved in regularly and systematically evaluating eLearning and eTeaching using a standard set of evaluative tools?

*Reporting & Evaluating***:** How often are key areas of eLearning and eTeaching systematically evaluated and reported?

*Modifications & Changes***:** What procedures are in place to ensure that appropriate changes are made to the course after the eLearning and eTeaching evaluations have taken place?

*Good Practice Dissemination***:** How do you integrate good practice messages into the eLearning course after they have been identify through the QA process?

*Complaints***:** In what ways are procedures for dealing with complaints from eTeachers and eLearners addressed?

Operational Management & Administration

- **Operational Management & Administration Factors:** What structures and procedures are in place for the management and administration of the eLearning provision?

*Roles & Responsibilities***:** How do you ensure that all those involved in the operational management and administration of the course understand their roles and responsibilities?

*Operational Management***:** What systems are in place to ensure the effective management of the course?

*Administration***:** What administrative support mechanisms are available for the eLearning and eTeaching course?

*Evaluation & Feedback***:** Do all members of the operational management and administration teams systematically evaluate their procedures and feedback to the central QA systems?

The standards on which the eQcheck is based are the Canadian Recommended eLearning Guidelines - the CanREGs, published and copyrighted by FuturEd Inc. and the Canadian Association for Community Education, 2002. The CanREGs are based on best practice and research in distributed learning and learning technologies, developed through a national consultation process, and sponsored by a number of national organizations. The key features of the CanREGs are that they are:

- consumer-oriented - developed with particular attention to return on investment in eLearning for learners

- consensus-based - developed through consultation with a balance of provider and consumer groups in Canada and beyond
- comprehensive - inclusive of all elements of the learning system: outcomes and outputs, processes and practices, inputs and resources
- futuristic - describing a preferred future rather than the present circumstances for design and delivery
- distinctively Canadian - reflecting the highest of Canadian values and learning priorities
- adaptable - best used for adult and post-secondary education and training, but adaptable to other levels of learning services
- flexible -- not all guidelines will apply in all circumstances

Taking all these eLearning quality approaches into account in this study we selected "Quality Assessment Rubric for eLearning Design" because this was the most comprehensive and the best fitting approach to our study. This quality approach is explained as follows.

Aspects for Quality Assessment

Learner Support and Resources

- Course Information
- On-line Support
- Content Support
- Channels for Feedback on Resources and Support

Online Organisation and Design

- Completeness
- Clear Syllabus
- Aesthetic Design
- Consistent and Functional
- Accessibility
- Channels for Feedback

Instructional Design and Delivery

- Opportunities for Interaction
- Alignment of Course Objectives
- Clearly Defined Learning Outcomes
- Variety of Learning Tasks
- Critical Thinking
- Channels for Feedback and Instructional Design

Assessment and Evaluation of Student Learning

- Opportunities for Self-Assessment
- Alignment between Objectives, Activities and Assessments
- Comprehensive Assessment Strategy
- Opportunities for Students to Receive Feedback
- Channels for Students Feedback on Assessment Strategy

In this approach they determine that the specific topics within each area are evaluated using a four-point scale:

1. Lacking – A given topic is not positively addressed. Development is needed.

2. Baseline – The quality of a given topic is positive but does not meet expectations, considerable improvement is needed.

3. Effective – The quality of a given topic meets expectations, however some improvement is needed.

4. Exemplary – The quality of a given topic exceeds expectations.

So depending on these scales we assumed that *exemplary* scale is the best eLearning quality target that has to be taken into consideration while designing an eLearning system. Therefore, while performing our checklist we used *exemplary* definitions for eLearning quality approach.

2.3 E-LEARNING USABILITY EVALUATION APPROACHES

2.3.1 Definition of Usability

Whoever has spent hours figuring out how to set a VCR clock already has a good idea of what usability is not. Usability can be simply defined as ease of use, the facility with which one can get something doing what it is intended to do. It can apply to practically any object that gets put to use for some purpose.(McNamara,2003)

Usability has been defined by some as the extent to which an application is learnable and allows users to accomplish specified goals efficiently, effectively, and with a high degree of satisfaction. An additional component that should be added to this definition is usefulness;

that is, a highly usable application will not be embraced by users if it fails to contain content that is relevant and meaningful to them(Miller,2005)

Usability is an approach to product development that incorporates direct user feedback throughout the development cycle in order to reduce costs and create products and tools that meet user needs.

Two international standards define usability and human-centred (or user-centred) design:

"Usability refers to the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of user." (ISO 9241-11)

"Human-centred design is characterised by: the active involvement of users and a clear understanding of user and task requirements; an appropriate allocation of function between users and technology; the iteration of design solutions; multi-disciplinary design." (ISO 13407)

There are many eloquent additional definitions of usability from books by usability professionals.

Usability means that the *people who use the product* can do so *quickly and easily* to accomplish *their own tasks.* This definition rests on four points: (1) Usability means focusing on users; (2) people use products to be productive; (3) users are busy people trying to accomplish tasks; and (4) users decide when a product is easy to use (Janice and Ginny, 1999).

After all, usability really just means that making sure that something works well: that a person of average (or even below average) ability and experience can use the thing - whether it's a Web site, a fighter jet, or a revolving door - for its intended purpose without getting hopelessly frustrated (Krug, 2000).

These different meanings can be described in four key requirements:

- **Usability means thinking about how and why people use a product.**

Good technical writing, like good interaction design, focuses on user's goals. The first step in creating a usable product understands those goals in the context of the user's environment, task or work flow, and letting these needs inform the design.

- **Usability means evaluation.**

Usability relies on user-feedback through evaluation rather than simply trusting the experience and expertise of the designer. Unlike conventional software acceptance testing, usability evaluation involves watching real people use a product (or prototype), and using what is learned to improve the product.

- **Usability means more than just "ease of use"**

The 5 Es – efficient, effective, engaging, error tolerant and easy to learn – describe the multi-faceted characteristics of usability. Interfaces are evaluated against the combination of these characteristics which best describe the user's requirements for success and satisfaction.

- **Usability means user-centered design**

Users are satisfied when an interface is user-centred – when their goals, mental models, tasks and requirements are all met. The combination of analysis, design and evaluation all approached starting from the user's point of view creates usable products.

The definition of usability in the (ISO 9241) standard is:

"The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use"

This definition can be expanded, and made more comprehensive, by including five characteristics which must be met for the users of a product:

- Effective
- Efficient
- Engaging
- Error Tolerant
- Easy to Learn

Effectiveness is the completeness and accuracy with which users achieve specified goals. It is determined by looking at whether the user's goals were met successfully and whether all work is correct.

It can sometimes be difficult to separate effectiveness from efficiency, but they are not the same. Efficiency is concerned primarily with how quickly a task can be completed, while effectiveness considers how well the work is done. Not all tasks require efficiency to be the first principle.

The quality of the user assistance built into the interface can have a strong impact on effectiveness. The effectiveness of an interface often relies on the presentation of choices in a way that is clearly understandable to the user. The more informative an interface can be, the better users are able to work in it without problems. Good interface terminology will be in the user's language and appropriate to the task.

Efficiency can be described as the speed (with accuracy) in which users can complete the tasks for which they use the product. (ISO 9241) defines efficiency as the total resources expended in a task. Efficiency metrics include the number of clicks or keystrokes required or the total 'time on task'

Navigation design elements such as keyboard shortcuts, menus, links and other buttons all have an impact on efficiency. When they are well-designed, with clearly expressed actions, less time and effort are needed for the user to make navigation and action choices.

Making the right choices for efficient use of the software depends on an understanding of the users and how they prefer to work. For example, are they likely to use the interface infrequently or to be habitual users who might learn hidden controls and shortcuts? Do they use the keyboard, mouse or other input devices? For example, keyboard shortcuts can be extremely efficient for proficient users who work with the interface intensively. If they are the primary interaction tool, they can slow down users who are unfamiliar with them, or with the software.

An interface is engaging if it is pleasant and satisfying to use. The visual design is the most obvious element of this characteristic. The style of the visual presentation, the number, functions and types of graphic images or colors (especially on web sites), and the use of any multimedia elements are all part of a user's immediate reaction. But more subtle aspects of the interface also affect how engaging it is. The design and readability of the text can change a user's relationship to the interface as can the way information is chunked for presentation. Equally important is the style of the interaction which might range from a game-like simulation to a simple menu-command system. (Bergman and Haitani, 2000)

Error tolerance: The ultimate goal is a system which has no errors. But, product developers are human, and computer systems far from perfect, so errors may occur. An error tolerant program is designed to prevent errors caused by the user's interaction, and to help the user in recovering from any errors that do occur.

Note that a highly usable interface might treat error messages as part of the interface, including not only a clear description of the problem, but also direct links to choices for a path to correct the problem. Errors might also occur because the designer did not predict the full range of ways that a user might interact with the program. For example, if a required element is missing simply presenting a way to fill in that data can make an error message look more like a wizard. If a choice is not made, it can be presented without any punitive language.

Some guidelines for preventing errors are:

- Make it difficult to take incorrect actions. Design links and buttons to be distinctive, use clear language, avoiding technical jargon, and be sure that dependent fields or choices appear together.
- Make it difficult to take invalid actions. Limit choices when possible to those which are correct, provide clear examples for data entry, present only appropriate navigation options.
- Make it difficult to take irreversible actions. Provide the ability to back track, provide means to undo or reverse actions, and avoid dead-end screens. Don't indiscriminately use confirmations – users become insensitive to them.
- Plan for the unexpected. Allow for users to add new entries, take exceptional routes through the interface or make choices you did not predict.

Easy to learn: One of the biggest objections to "usability" comes from people who fear that it will be used to create products with a low barrier to entry, but which are not powerful enough for long, sustained use.

But learning goes on for the life of the use of a product. Users may require access to new functionality, expand their scope of work, explore new options or change their own workflow or process. These changes might be instigated by external changes in the environment, or might be the result of exploration within the interface.

An interface which is easy to learn allows users to build on their knowledge without deliberate effort. This goes beyond a general helpfulness to include built-in instruction for difficult or advanced tasks, access to just-in-time training elements, connections to domain knowledge bases which are critical to effective use. (STC Conference 48., 2001)

User Centred System Design is based on understanding the domain of work or play in which people re engaged and in which they interact with computers, and programming computers to facilitate human action.

User-centred design (UCD) is an approach to design that grounds the process in information about the people who will use the product. UCD processes focus on users through the planning, design and development of a product.

There is an international standard that is the basis for many UCD methodologies. This standard (ISO 13407): Human-centred design process) defines a general process for including human-centred activities throughout a development life-cycle, but does not specify exact methods. This standard can be seen in the figure 2.3.

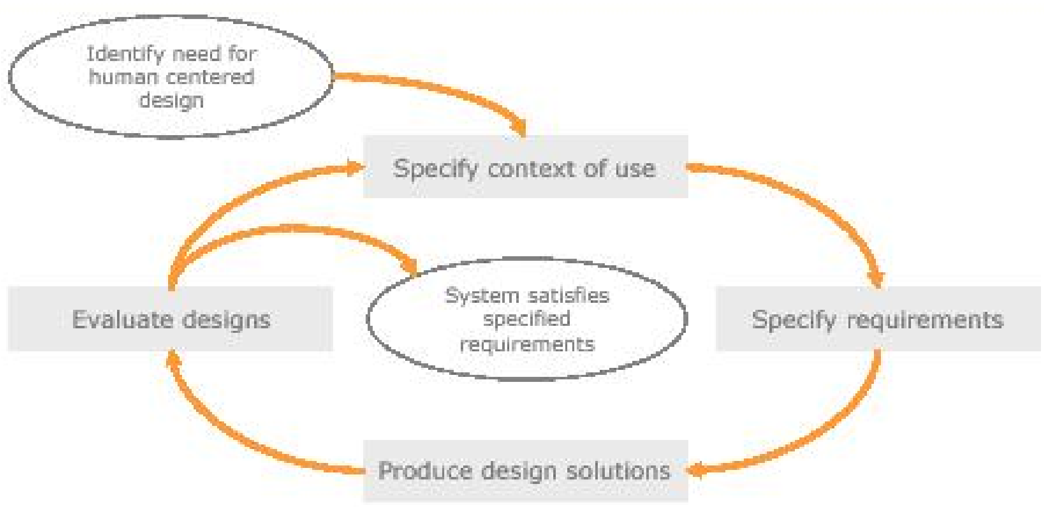

Figure 2.3 – Human-Centred Design (adapted from ISO 13407)

In figure 2.3, once the need to use a human centred design process has been identified, four activities form the main cycle of work:

1. Specify the context of use**:** Identify the people who will use the product, what they will use it for, and under what conditions they will use it.
2. Specify requirements**:** Identify any business requirements or user goals that must be met for the product to be successful.
3. Create design solutions**:** This part of the process may be done in stages, building from a rough concept to a complete design.
4. Evaluate designs**:** The most important part of this process is that evaluation - ideally through usability testing with actual users - is as integral as quality testing is to good software development.

The process ends - and the product can be released - once the requirements are met.

Heuristic evaluation is an informal evaluation method proposed by (Nielsen,1994), whereby experts – guided by a set of heuristics or usability principles – evaluate designs, storyboards, or operational systems. Nielsen's classic set of ten usability heuristics comprises:

- visibility of system status,
- match between system and real world,
- user control and freedom,
- consistency and standards;
- error prevention;
- recognition rather than recall,
- flexibility and efficiency of use,
- aesthetic and minimalist design,
- help users recognize, diagnose, and recover from errors,
- help and documentation.

By filtering these heuristics through the concepts of *cognitive authenticity*, *contextual authenticity* and *socio-constructivism*, (Squires and Preece, 1999). They customized them to systems for instruction and learning, to be used in predictive evaluation prior to selecting systems for use. In the study of (Villiers,2004) they were converted to evaluation questions for a questionnaire survey, undertaken among learners *after* use. He uses a 4-point semantic scale as *Strongly Agree, Agree, Maybe and Disagree.* The relevant questions of his checklist approach are given after each related dimension:

1. Match between designer & learner models: This is determined by considering intrinsic feedback and whether it represents cognitive tasks in ways that foster formation of a learner model consistent with the designer model.

- Doing exercises helps me understand the theory of relations.
- The feedback (system responses) to my incorrect answers helped me to get them right.
- Some of the exercises were really complicated.

2. *Navigational fidelity*: Investigating navigational fidelity involves considering the structure of navigation, cosmetic authenticity and the effectiveness of limited representation of the world, as provided by a system.

- When I use the eLearning system, I know exactly where I am, what parts I have done, and what I should still do.
- The lesson is easy to operate.
- I got ‘lost’.

3. *Appropriate levels of learner control*: This relates to the balance between learner control, self-direction, customization, consistent protocols, and system responsibility.

- I enjoy being able to choose what to do next.
- I appreciate being able to choose whether to do the theory and the exercises or go straight to the exercises.
- I would like to choose the level of difficulty of the exercises.

4. *Prevention of peripheral cognitive errors*: There is a relationship between domain complexity and error prevention. While usability-related errors should be avoided, cognitive errors are part of the learning process, substantiated by Mayes and Fowler [1999:485] who stress that in educational applications a ‘seamless fluency of use is not necessarily conducive to deep learning the software must make learners think’. Peripheral usability errors should be anticipated and avoided and where possible, novice versions could be provided.

- I made mistakes in the way I use the eLearning system, i.e. mistakes in using the system.
- I made mistakes in doing the exercises.
- I made mistakes because I used operations and keystrokes I know from another system.

5. *Understandable and meaningful symbolic representation*: Representational forms and symbols should be considered. Interfaces should present low cognitive demands and learners should not have to remember the forms of interaction. The symbols, icons and names used for learning objects should be those of the subject domain and should be used consistently.

- The symbols and names that represent mathematical objects are used consistently in the lesson.
- It is easy to grasp the structure of the screen displays.
- eLearning system helps me understand a difficult section of the courses.
- I got 'stuck'

6. *Support personally meaningful forms of learning*: There are multiple representations, moreover the software will be used in tandem with various learner support materials. Metacognition should be supported and the software should indicate clearly what learning styles are supported.

- When a concept was taught or illustrated in more than one way, it helped me understand it.
- I enjoy approaching studies collaboratively, i.e. working with a fellow-student.
- I like doing examples on the computer.

7. *Strategies for cognitive error recognition, diagnosis and recovery*: Established techniques, such as cognitive conflict, scaffolding, and bridging should be used to promote the recognition–diagnosis–recovery cycle.

- The feedback to my incorrect answers was useful.
- When I got an answer wrong, I was able to get it right on the next try.
- There is too much information on the screens; it confuses me.

8. *Match with the curriculum*: Software should correspond to curricula and lend itself to educator-customisation.

9. *Distinctive features*, unique to the environment being evaluated.

10.*Capacity of the system* to engage learners and hold their attention(Villiers,2004).

2.3.2 Ways to Make eLearning More Usable

Many sources of usability advice, both printed and on the web, are readily available either as a checklist of items or discussed at greater length. It is impossible to provide a definitive list of concrete steps to take that are applicable to every case; each learning situations's specific goals determine the questions used to judge whether learners have achieved them. In addition to basic interface factors, specific questions about what has been learned and how it was learned would need to be formulated for each unit of content. Because general usability guidelines do apply as well to eLearning content, first some overall usability considerations will be presented below. Then several authors' advice about improving eLearning interfaces in particular will follow. Finally, sources of several usability checklists available for use with educational content will be provided (McNamara, 2003).

General Usability Guidelines

The following universally applicable usability tips are taken from Steve Krug's "Don't Make Me Think - A Common Sense Approach to Web Usability". (Krug, 2000)

- Don't make me think. The main title of the book is his first rule of usability. This applies to elements that are not core content (navigation, layout, etc.) which should be self-evident, obvious and self-explanatory.
- Provide a clear visual hierarchy. This is crucial because we don't really read computer screens, we scan them. Especially, users rarely read instructions.
- Design a simple and quiet layout with clearly defined areas, using conventional elements. Conventions become conventional because they work and are therefore commonly accepted.
- Make clickable items obvious and the result of clicking a button an unambiguous choice. Time should not be lost figuring out whether a buton is a button, and what happens if one clicks on it.

- Minimize text. Krug's advice is actually to remove 50% of the text, then remove half of what's left.
- Provide navigation that substitutes for instructions. Well designed navigation should be all that a user needs to know where to begin and what her options are.
- Use vivid, saturated colors. One out of 200 women and one out of 12 men have trouble detecting some color distinctions.
- Provide a clear starting point. Whether the path is linear or in the form of a mind map, starting point(s) should be obvious.
- Path type should be obvious. The user should be able to tell if she is expected to follow a linear path or a flexible path of her choice.
- Use Flash wisely. Content in Flash format should be included only if it is presented in small well thought out units that add useful functionality.

eLearning-specific Usability Guidelines

Jakob Nielsen confirms that general usability principles apply to eLearning, and adds the following advice specific to eLearning (Nielsen, 2001).

- Pay attention to response time. Since it is necessary to keep content fresh in the learner's mind, it is all the more important not to force the learner to wait for slow downloads or other delays.
- Provide experience based learning. Nielsen suggests using computers to offer that which a good book cannot: simulation systems, problem based learning, case studies or calculation exercises. Allow learners to try out and discover things themselves in well planned situations.
- Make discussion forums easy to understand. Though this form of interaction among students can offer the contact that so enhances learning, difficult interfaces for discussion groups may prevent some students from even discovering the option. Kevin Kruse organizes usability advice for eproducers into 4 categories: assist user memory, put the user in control, design logical and consistent screen elements and provide user guidance (Kruse, 1999).

From Kruse's first category, assisting user control:

- Chunk information into meaningful blocks, organize the menu structure. Some specifics include limiting menus to 7 items and matching the placement order of menu items with the structure of tasks.
- Use mental models or visual metaphors. A mental model should contain no more than 3 layers or paths.
- Don't overload the sensory system. Too much stimulation can quickly overload the human sensory system.
- Use multiple access points. Provide several ways to get to content, for example, via the main menu, bookmarks, an index, the search option and a site map.

From Kruse's second category, put the user in control:

- Provide status messages. If the computer is busy doing something, a status message gives the user the impression of being in control.
- Allow reversible actions. An undo or back button conveys the idea of forgiving software, reducing the learner's anxiety and increasing his confidence.
- Allow both mouse and keyboard input. For different types of users, the ability to use either the mouse or the keyboard increases productivity as well as accessibility.
- Provide access to Help, Menus and Exit with one click. These are frequently used functions that require instant access.

From Kruse's third category, logical and consistent screen design:

- Use logical screen layouts. This includes taking reading patterns into account, such as western readers' Z path across and down the screen.
- Be consistent with media choices. The use of audio narration or video based feedback should be provided consistently after being introduced so students don't wait in vain for something that doesn't happen.
- Have menus behave predictably. If menu items open a submenu or lead to a learning activity, they should do only that and not mix action types.

From Kruse's last category, provide user guidance:

- Include page counters. Learners need to know which page they're on and the total number of pages.

- Give appropriate warnings. Before sending log in data, taking a test or exiting a program, users should be presented with a confirmation screen.

Dave Smulders' article (Smulders) on designing for learner-users gives one specific tip that cannot be forgotten:

- Know your users. As mentioned above, this includes cultural and technical aspects as well as where they stand academically.

Based on tests conducted by Frontend (Frontend, 2001), two additional tips emerge:

- Allow comparison tasks. For example, using mouse-over events to display item descriptions prevents viewing two descriptions at once.
- Allow resizable text. Most browsers let the user increase text size – this has become expected and should be incorporated into eLearning programs for accessibility reasons also.

In his article on usable eLearning, Michael Feldstein wisely emphasizes the cognitive goals of eLearning content. His suggestions are not actually usability tips, but rather a series of questions that must be addressed by research into this area of eLearning. They are presented here as such, as they may also serve as a basis for reflecting on usability for eLearning.

- To what extent do navigation elements help learners internalize and remember the structure of the content?
- To what extent do navigation elements help learners find key concepts for later review?
- Does audio narration doubling of text presentation affect the learner's ability to remember key facts and concepts, or to process complex concepts?
- Does a threaded discussion board interface affect the frequency with which learners collect particular ideas or facts?
- Does a threaded discussion board interface affect the frequency with which learners synthesize various viewpoints in a conversation?

2.3.3 Application of Usability Techniques to eLearning

Defining eLearning in a very general sense as any learning situation that happens with the help of a computer, then situations where usability come into play are the those where someone sits in front of a computer to interact with learning material. The obvious case is when the learner herself starts an online course and begins to execute the steps required to "take the course". In this case the course material itself would be that which should be usable. Not necessarily the course itself, nor the individual modules it is made up of, but the individual "pages" or screens constituting a module are the smallest units where one would begin to apply usability principles. They would of course be applied consistently over the entire course (McNamara, 2003).

A step up from a single online course one may encounter a Learning management system (LMS), offering the administrative functions related to taking and offering online courses. If such a system is present, the student is required to log on to it before accessing course material. Here too is an eLearning situation where usability cannot be forgotten.

Other categories of users also come into contact with an LMS: teachers, system administrators, course developers and designers. Considering the variety of tasks that all these users need to accomplish, ease of use of the LMS is again essential. Unfortunately, this type of eLearning software is often difficult to use. (Piguet and Pareya, 2000)

Changing scale from a detailed to a more general level of course organisation, one may ask whether usability can be applied at the course design level. If one admits that the user interface is a reflection of overall course design (main menu items representing course chapters, for example) the dividing line between interface design and course design begins to blur. Questions concerning user interface usability, such as "does the user easily understand which buttons to click in order to complete a certain task?" have parallels where course design is concerned, for example: "does the learner quickly understand what a certain task will help her learn?". Referred to as "learnability", this is indeed the principle goal of any online educational material (McNamara, 2003).

2.3.4 Need of Usable eLearning Systems

Indeed, if a learner is quickly confused about how to navigate around course material to the extent that he's not able to take advantage of whatever learning opportunity he might have found, the chance to target the intended educational goal is missed. Imagine a scene by a serene lake in the mountains: here, we equate acquiring knowledge with the ability to sit calmly by the water and absorb the sights, sounds, smells and textures of the surroundings – this is the learning place. In contrast, imagine a different scene in a traffic jam in the city: sitting in a cramped car, frustrated by confusing road signs and a bad map, distracted by noise and stressed by uncertainty. This learner, trying to find his way to the mountain lake, the learning place, is blocked by the obstacle of poor usability - no learning can happen.

Dave Smulders outlines the need to consider the eLearning consumer as a kind of double persona, the learner-user. A distinction is made between form, the user interface, and content, the learning material, where the user part of the persona is concerned with the form and the learner is interested in the content. "By navigating their way through the form of your Web-based environment, users can access the content, at which point they can don their learner's cap..." (Smulders) The learner-user operates differently depending on which persona is called upon: "using" an eLearning site means making functional connections, whereas "learning" at that site means making cognitive connections. Clearly, doing this cognitive work of learning requires mental energy on the part of the learner. The thinking budget is better spent on the content than wasted on the navigation, as suggested by Steve Krug's "first law of usability: don't make me think!". (Krug, 2000) If making functional connections can be cognitively effortless, those cognitive resources are made available for the real work: learning.

In a learning situation where usability has not been made a priority, the resulting frustration can only dampen a student's motivation. Given that motivation is a significant factor in determining the success of any Learning endeavor, creating usable eLearning can avoid the drop in motivation caused by a loss of orientation. As Kevin Kruse puts it, "the success of any training program is largely dependent on the student's own motivation and attitude. If a poorly designed interface has them feeling lost, confused or frustrated, it will become a barrier to effective learning and information retention." (Kruse) Jakob Nielsen points out that it is easy to lag behind in an online course or be distracted by other more immediate cicumstances, reducing motivation to continue. "Online courses are inherently not very motivational and not as effective as traditional courses. Hence, there is all the more reason for you to keep the user experience good and engaging..." (Nielsen, 2001).

There are also other practical reasons why designing usable eLearning makes sense. A negative user experience can have an effect on the long-term success of eLearning programs by increasing the number of unsatisfied learners. A student may complete a course she signed up for in spite of her dissatisfaction with the interface, but is very unlikely to sign up for another course afterwards. Referring again to his double learner-user persona, Smulders states, "Online courses that are designed for learners without any thought to users invariably results in frustrated students", and further, "...as online courses from institutes of higher education have flooded the marketplace and become a constant source of criticism for their poor design and high drop-out rates, it is clear that the rarely made distinction between learners and users is problematic." (Smulders).

2.3.5 Integration of Usability into eLearning

Usability experts generally agree on the answer to this question: test early and test often. Even with attempts to incorporate usability guidelines into eLearning design, testing is still highly recommended. The reason for this is that the people involved in designing online courses, or any product to be used by others, are themselves already too familiar with how it is intended to work and lack the distance necessary to pinpoint problem areas. Only test users who have not yet had any contact with the product can give an accurate idea of how easy or difficult it is to use (McNamara, 2003).

Usability testing has not as yet become a standard step in the online course development process. A common assumption is that usability testing is an expensive undertaking carried out only by specialized groups of experts. Though this can be true, it need not be the case, as will be shown by other less expensive options mentioned below (McNamara, 2003).

The point of testing for usability is to expose prototypes to test users unfamiliar with it. Their experiences are valuable and essential to finding out just how well a program "works" in very practical terms. They can address questions such as: Is it easy to find one's way through an activity or site? Are there places where the user didn't understand where they were or what they could do? or, Is the organisation and purpose of the material clear? Testing for learnability could be applied as well by asking slightly different questions of the test learners, for example: Was the goal of a certain learning activity evident? What is being taught on the site? How should a learner procede through the course material? Michael Feldstein insists on the importance of eLearning specific testing: "If we are serious about making our eLearning usable, we in the field must make a concerted effort to define usabilityquestions that are related to the learners' cognitive goals..." (Feldstein, 2002).

There are several ways to arrange user testing. In an ideal situation with a substantial budget, one would hire usability experts to set up an optimal test situation with large groups of carefully selected test users. If funds are available, then it certainly makes sense to opt for high budget testing; if not, however, simpler and still quite effective options are available that provide very reliable results. The basic model is common to most approaches: a test user is faced with a program and left to her own devices to figure out how it works. She will be asked questions about her experience and encouraged to think aloud during the testing to give as much feedback as possible about how quick and easy (or not) the program is to use. Some of the more costly testing scenarios may include hand picking a large group of test users who represent the profile of the intended user audience. Testing sessions can be recorded for viewing synchronously or after the session. Special hardware and software may track the user's eye movements to record the visual paths that they followed. Mouse movements can be recorded and later analyzed by screen recording software.

Some of these options may only be affordable in a research setting however. In the day to day work of producing online educational courses it is more common that financing is not abundant, yet this does not preclude carrying out viable user testing. Steve Krug's highly recommendable book, "Don't Make Me Think" (Krug, 2000) provides excellent advice on how to conduct usability testing at low cost and still obtain high returns. The subtitle of the chaper on testing emphasizes a crucial point: "keeping testing simple - so you do enough of it". His advice is to do simple testing early, and repeat the process as often as possible. The simplest approach could consist of just sitting next to individual test persons while each one makes his way through the learning material, taking notes and asking questions where appropriate. Because the goal of testing is to inform - not, as is often the belief, to do a scientific experiment - a test with one single user will give valid results. Showing that 2 tests with 3 users reveals more problems than 1 test with 8 users, Krug insists on conducting several simple tests during development.

Heuristic testing is another approach that doesn't require recruiting real test persons. Here a small team of experts would be asked to examine online content for usability. Another proponent of low cost usability testing, Michael Feldstein calls for research in order to determine best practices specific to eLearning. These would then be applied by experts during heuristic testing. (Feldstein, 2002).

Paper prototyping is another even lower cost option. Applied very early in the development phase of a project, this testing is carried out using simple paper printouts of storyboard screens. (Nielsen, 2003) Significant development cost can be saved for two reasons. First, no time is lost developing a prototype that will require modification after the paper prototypes are tested. Second, user feedback is incorporated into development at the earliest possible stage of development, providing essential information that eliminates development time that would have been spent modifiying initial working prototypes.

Developing for so-called personas, a process incorporating usability during the entire design period, is an approach suggested by usability expert Alan Cooper. Rather than testing on real people, the project team invents a detailed persona closely representing a typical user for their site. Any design questions that emerge during development would be addressed specifically with this persona in mind. An effort is made to create as realistic a virtual test person as possible, including minute details like name, job, family, hobbies as well as their computer habits, whom they can "question" about their preferences.

Every university campus is teeming with potential test users: students. Attracting students to become test users doesn't need to be difficult. With fair monetary compensation for an hour or two of work, short-term commitment to abide class scheduling and the extra satisfaction of contributing to actual course development can be convincing factors. It's imaginable that a pool of willing and available testers could be formed during the academic term, and later called upon as needed during development.

Whatever the type of testing that is decided upon, it is important to remember that the process is iterative. Each round of testing will reveal the obvious points that somehow got overlooked. After testing the project goes back to development to resolve the problems uncovered in the test. Then another round of testing should be scheduled, followed by further development, and so on. Each testing session helps to clarify uncertainties encountered by the learner, allowing the team to produce ever more effective eLearning materials (McNamara, 2003).

2.3.6 Factors Affecting eLearning Usability

Following usability principles does not consist of simply following a fixed set of rules. Where learning content is concerned, the specific cognitive goals of a single module or exercise would determine the questions used to judge the usabilty or learnability of the content. Usability is a dynamic concept, evolving and differing over time and place (McNamara, 2003).

The characteristics of any software that are perceived as adding to it's ease of use are in fact the product of our collective habits over time. What we are used to is what we understand easily and quickly, it is what we have come to expect and is therefore intuitive to us. Habits change over time however, so that what we are used to today does not resemble what we were used to in the past. Computer technology and interfaces change quickly - what was new and strange to us just months ago may today seem quite familiar. As an example, consider the use of tabs (that mimic paper file folder tabs) as navigational interface elements. Not long ago they were novel, and perhaps took a little time to get used to; today they are common features of user interfaces. The look of buttons has evolved as well, where today we can recognize more than one visual representation of the places where we can expect to be able to click and make something happen. As Colin Moock puts it so well, "'usable' isn't inherent in a technology, it's a correspondence between a tool and the habits of the person that uses the tool." (Moock, 2003).

Not only is the concept of what is usable dynamic, differing from one time period to the next, it is also variable depending on cultural context. This relates to a rule that is basic to designing usable online courses or any online content: know your user. This means that it is indispensible that that cultural context of the expected users be taken into account. A difficult challenge of those producing eLearning content is how to create a positive user and learner experience for different cultures. Issues that need to be addressed include language, reading habits, visual metaphors, age, etc. Dealing with language differences often involves more than simple translations, as learners from different language groups come to an online course with different histories and perspectives that go beyond simple choice of language (McNamara, 2003).

Reading habits are also an expression of our cultural environments. A clear example of this is the assumption about how western cultures read in a Z pattern, left to right and top to bottom. (Kruse) Based on this, design for western users would incorporate that assumption into decisions about content placement on a page. In cultures where text is read from right to left, however, not only would the language need to be adapted but the relative location of the text on the screen.

Visual metaphors are often used as navigational aids. In addition, graphical content can play a major role in eLearning content due to it's powerful effect on information retention. What one associates with any given visual representation is strongly related to cultural heritage; choosing an inappropriate image for a target culture could at best reduce Learning effectiveness - at worst, it may actually insult the learner (McNamara, 2003).

When following the "know your user" rule, the user's cultural background is just the beginning of the picture. In addition to recognizing who they are, it is necessary to find out what equipment they will be using, the speed of their internet connection and their computer skills. In order to avoid forcing the user to wait while content downloads, it might be necessary to eliminate heavy image files, for example, or to offer more than one option and let the user decide which to view. Dave Smulders offers an excellent example with a case of promoting collaborative work among students. Depending on the students' technological surroundings and experience, the necessary activities might be well suited or completely off base. He describes possible scenarios thus: "are they spread across a city or state in three or four high-tech labs with video conferencing facilities, or are they employees of an international non-profit agency with field offices in multiple time zones and equipped with unreliable technology?" (Smulders) It would also be necessary to know whether the users are familiar with the software and equipment or if the experience would be entirely new. Knowing the difference in the design phase makes all the difference during the learning phase.

2.3.7 Methods for eLearning Usability Evaluation

Various usability evaluation techniques exist; choosing among them is a trade-off between cost and effectiveness. Some methods, like heuristic evaluation, are easier to administer and less costly, but there are problems with using such methods. These problems mostly come from applying a small set of principles, the heuristics, to a wide range of systems. In (Notess, 2001), it is asserted that usability testing needs additional consideration in the light of the web-based learning environments. A number of proposals try to meet this requirement. (Dringus, 1995) proposes that usability heuristics summarized by (Shneiderman,1997) and (Nielsen, 1993) can be applied to evaluate eLearning applications interfaces as well (Dringus, 1995). Ravden and Johnson provide a checklist that emphasizes on visual clarity, consistency, appropriate functionalities, flexibility and control, and so on (Ravden and Johnson, 1989). Schwier and Misanchunk present principles of simplicity, consistency, clarity, appropriate use of white space, time and minimal memory load (Schwier and Misanchunk, 1993). In (Squires and Preece, 1999) it is proposed that an adaptation of Nielsen's heuristics taking into account socio-constructivism tenets, i.e. match between designer and learner models, support for personally significant approaches to learning, match with the curriculum, appropriate level of learner control. However, these proposals yet validated with user studies(Costabile et al., 2005).

There are many ways to determine the usability of an eLearning application, including heuristic evaluations, usability tests, and field studies.

Heuristic evaluations: A heuristic is a rule or well-established standard. A heuristic evaluation is a technique that entails the formal review of an application with experts in usability and interface design to determine whether the application is aligned with recognized and established standards for graphical user interfaces. Ideally, this procedure will be conducted before the application goes live or reaches another stage of development. The primary goal of a heuristic evaluation is to identify potential usability and ease of use issues in order to resolve them before final implementation.

Usability tests ask users to perform specified tasks on an application within a controlled laboratory environment. Typical metrics collected during usability tests include the levels of success users have performing a task, the amount of time that users need to complete a particular **task, and the level of satisfaction that users have with the application.**

Field studies involve watching users interact with the application in their own environments. The appeal of field studies is that they negate the need to make assumptions concerning how learners will use the application. Further, they frequently provide additional insight into the wants, needs, and expectations of learners.

In truth, there are many challenges associated with evaluating the usability of eLearning applications. For example, one challenge for eLearning applications is that they must accommodate the diverse backgrounds, experiences, and learning styles of users.

At this point, it's worth noting that determining the usability of an application is not the same as evaluating its educational effectiveness. That is to say that a highly usable product is no guarantee that learners will retain information. Nonetheless, it is virtually assured that an application with low levels of usability will not enable learners to access and assimilate information at all. (Miller,2005)

(Ardito et al., 2004) handles eLearning usability evaluation in a different aspect of view. They claim that an eLearning platform (container) is an environment integrating tools and services. Attributes for a platform generally differ from those of a specific eLearning module (content), since different features must be considered. However, some characteristics of the content provided through a platform are bound to functionalities of the platform itself. In identifying criteria and attributes for evaluating eLearning tools, we must consider nature and goals of eLearning, requiring user-system easy interaction, but even significant knowledge gain.

2.4 COMPARISON OF QUALITY EVALUTION APPROACHES AND USABILILTY EVALUATION APPROACHES

It has been the major debate whether usability influences quality or the opposite. The objective of usability is to achieve quality of use. Usability requirements should be stated in terms of the effectiveness, efficiency and satisfaction required in different contexts. User-based evaluation can be used to validate achievement of these requirements. Usability attributes provide a contribution to achieving quality of use. The presence or absence of these attributes can be verified early in design. (Bevan, 1995) claims that quality of use should be the major design objective for an interactive product. This idea is presented in figure 2.4. This relates usability to business objectives and elevates usability from an optional extra to the prime design goal. The purpose of designing an interactive system is to meet the needs of users: to provide quality of use (see the figure below, adapted from the working draft of ISO/IEC 14598-1: Evaluation of Software Products). The users' needs can be expressed as a set of requirements for the behaviour of the product in use (for a software product, the behaviour of the software when it is executed). These requirements will depend on the characteristics of each part of the overall system including hardware, software and users. The requirements should be expressed as metrics which can be measured when the system is used in its intended context, for instance by measures of effectiveness, efficiency and satisfaction. At this level, the required system characteristics could be minimum values for the effectiveness, efficiency and satisfaction with which specified users can achieve specified goals in specified environments. The required values of these external metrics provide goals for design. To achieve these goals the internal attributes of the system can be specified as internal requirements. At this level usability requirements may be in terms of general principles (e.g. provide consistency, support the user's task), specific interface details (e.g. icons and menu design), or use of style guides. These attributes of the software can be evaluated to produce internal metrics verifying how closely the internal requirements have been met. Although these attributes contribute to achieving quality of use, users and tasks vary so much that no set of interface guidelines alone can ensure that a product will be usable.

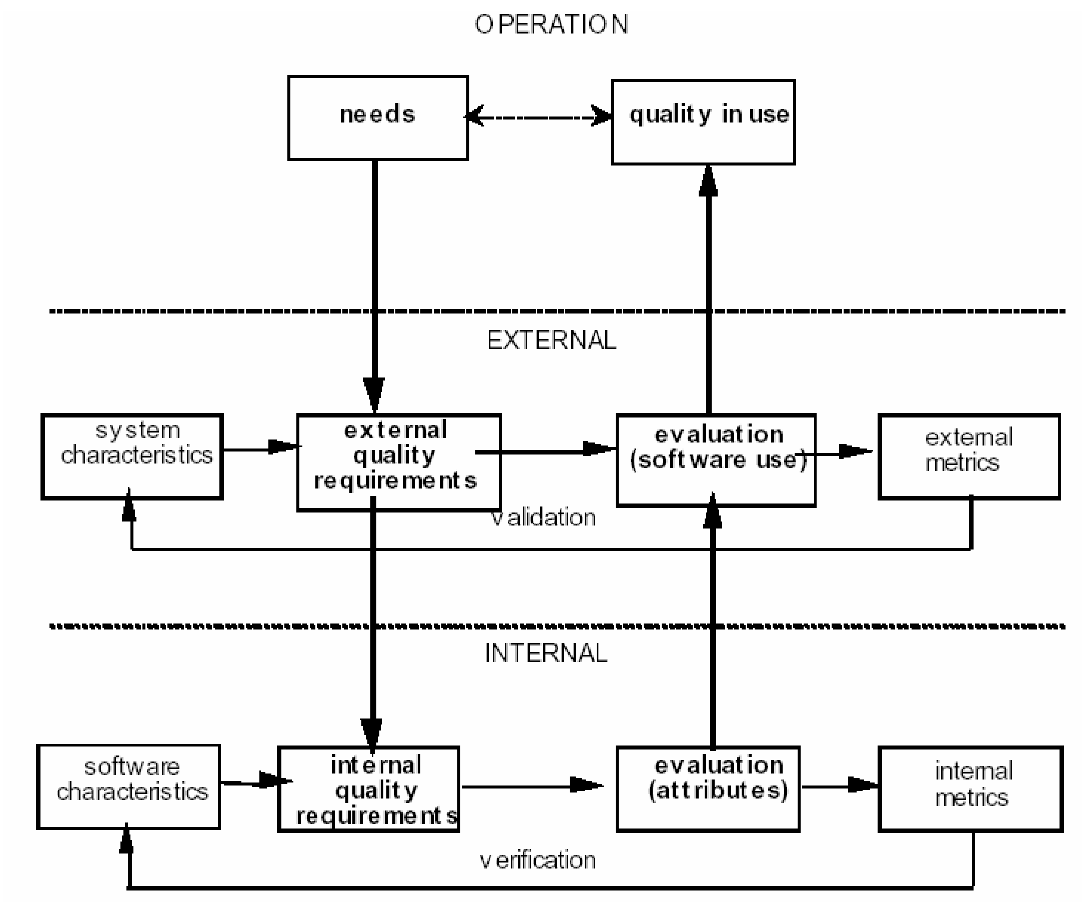

Figure 2.4 Quality Requirements in Design (adapted from Bevan, 1995)

The quality of use is determined not only by the product, but also by the context in which it is used: the particular users, tasks and environments as shown in figure 2.5. The quality of use (measured as effectiveness, efficiency and satisfaction) is a result of the interaction between the user and product while carrying out a task in a technical, physical, social and organisational environment (see the figure below, from Bevan, 1995). This means that there is no such thing as a "usable product" or "unusable product". For instance a product which is unusable by inexperienced users may be quite usable by trained users.

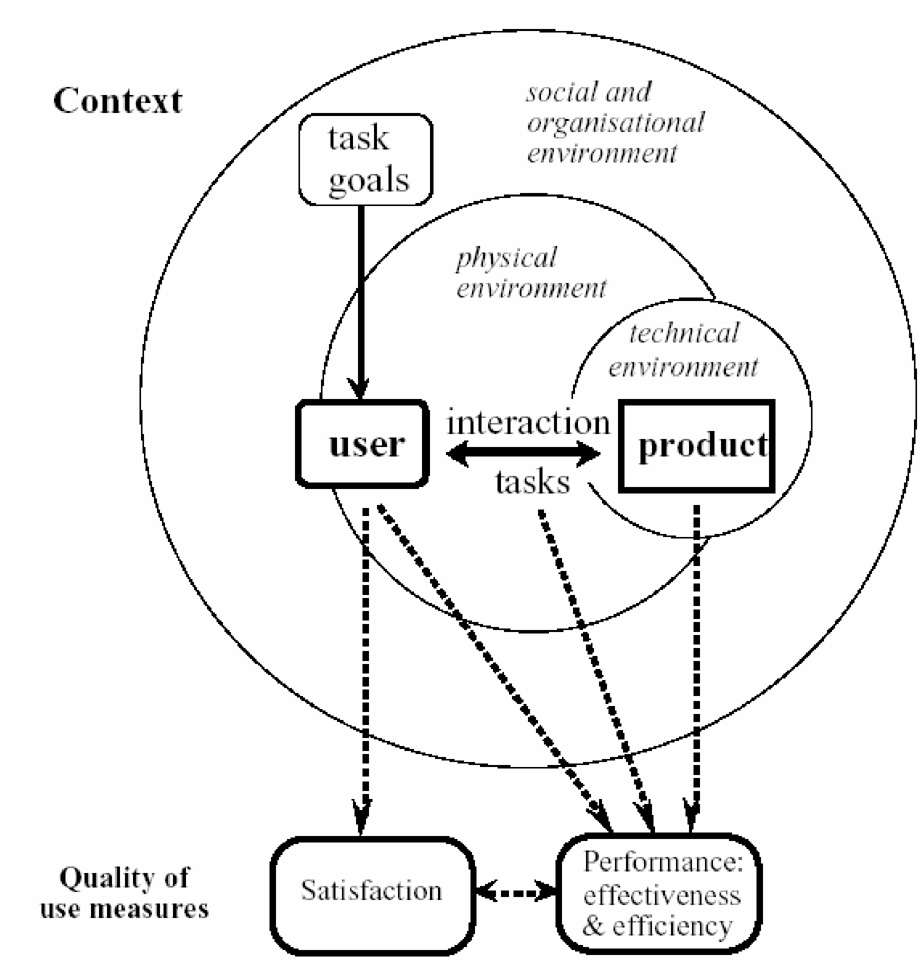

Figure 2.5 Quality of Use Measures Determined by the Context of Use (adapted from Bevan, 1995)

It is therefore essential to identify the intended context of use before carrying out any usability evaluation. In many cases it will be necessary to evaluate a product separately for different user groups carrying out different tasks. This applies both for evaluation of usability attributes and for evaluation of quality of use. For instance it may be necessary to consider different user groups when evaluating the appropriateness of the design and content of a help system. Similarly, when evaluating quality of use by user testing, it may be necessary to decide which combinations of user and task should be selected for evaluation. Depending on these information, we can summarize the relationship between usability and quality as the following: quality approaches the evaluation process in terms of some criteria but usability approaches the same evaluation process both the criteria and the performance criteria which are named as effectiveness, efficiency and satisfaction.

2.5 SUMMARY OF RESULTS OF CHAPTER 2

In chapter 2, traditional learning and eLearning are compared and the benefits of eLearning are defined. Usability evaluation and quality evaluation approaches of eLearning systems are compared. There are many overlapping items in these approaches depending on the natural result of interaction between usability and quality. There is evident need of developing a widely accepted quantitative method applicable to eLearning systems measuring the usability of eLearning systems and determining the usability problems.

CHAPTER 3

USELEARN: AN APPROACH FOR ADAPTIVE USABILITY EVALUATION OF E-LEARNING SYSTEMS

3.1 ELEARNING USABILITY CHECKLIST

The first step to create an eLearning checklist was to gather as many checklist approaches as we could. All the checklists that we could reach is shown in the table 3.1. These approaches were only the ones which are dealing with eLearning systems in the aspect of pure usability only. All these checklists were developed to evaluate the usability of eLearning systems for sure. But the names of the dimensions they use were different. We clustered all of them as shown in the table 3.1. Then we chose the most comprehensive one (Dringus et al., 1995 checklist approach) which had 13 dimensions to compare with the checklist approaches which are dealing with eLearning systems in the aspect of pure quality only.

Table 3.1 eLearning Checklist Approaches Dealing with Pure Usability

Understandable and meaningful symbolic representation	Visibility	Presentation	Site Design
Distinctive features,unique to the environment being evaluated Capacity of the system to engage learners and hold their attention	Aesthetics		
Capacity of the system to engage learners and hold their attention Match between learner and designer models	Course Management	Application Proactivity	Module Design
Capacity of the system to engage learners and hold their attention Match with the curriculum Support personally meaningful forms of learning	Interactivity		
Capacity of the system to engage learners and hold their attention Appropriate learner control	Flexibility		
Support personally meaningful forms of learning Match between learner and designer models Understandable and meaningful symbolic representation	Consistency		
Capacity of the system to engage learners and hold their attention	Reducing Redundancy	User's Activity	Content Organization
Prevention of peripheral cognitive errors Cognitive error recognition:Diagnosis and recovery	Error Prevention		
Understandable and meaningful symbolic representation	Functionality		
Match between learner and designer models Navigational fidelity	Feedback&Help		Navigation
Navigational fidelity	Memorability		
Match between learner and designer models Navigational fidelity	Efficiency		
	Accessibility	Hypermediality	
(Villiers, 2004)	(Dringus et al., 2005)	(Ardito et al, 2004)	(Miller, 2002)

Similarly, we searched the checklist approaches dealing with eLearning systems in the aspect of quality. Depending on our brainstorm sessions, we decided to use as a checklist tool “Quality Assessment Rubric for eLearning Design” which was again the most extensive one among the quality checklist approaches. Before getting this checklist as a whole we deleted some dimensions of it which seemed to be irrelevant with our case study. Then we tried to find out the overlapping and differentiating dimensions and gained a comparison of these checklist approaches which is related to usability and quality at the same time as shown in table 3.2.

Table 3.2 Comparison of the eLearning Usability-Quality Evaluation Checklist Approaches

Pure Usability Dimensions	Common Dimensions	Pure Quality Dimensions
Visibility	*Aesthetics*-Aesthetic Design	Clear Syllabus
Memorability	*Course Management*-Course Information, Online Support,Content Support	Alignment of Course Objectives
Flexibility		Clearly Defined Learning Outcomes
Reducing Redundancy	*Interactivity*-Opporunities for Interaction	Variety of Learning Tasks
Error Prevention	*Consistency&Functionality*-Consistent and Functional	Critical Thinking
		Opportunities for Self-Assessment
	Feedback&Help-Opportunities for Students to receive Feedback	Alignment between Objectives, Activities and Assessments
	Efficiency-Completeness	Comprehensive Assessment Strategy
	Accessibility-Accessibility	

There were nearly 200 questions in total including the usability and quality checklist. Apparently, it was impossible to give the test participants out a checklist containing so many questions and request them to evaluate it. So we tried to decrease the number of questions by selecting the ones that were measuring the concerning dimension of them most effectively. While creating this modified checklist approach, we merged the overlapping dimensions by trying to balance them with equal number of questions taken from both approaches if applicable. Here we intended to let the checklist fairly measure the usability and quality of the eLearning system at the same time. Of course, we had to rename the overlapping dimensions taken from both approaches in different names or choose one of the existing names. For example, for *Aesthetics* from usability approach and *Aesthetic Design* from quality approach we chose “*Aesthetics*”. And we named merged dimension as “*Completeness*” which is constituted by *Efficiency* coming from usability approach and *Completeness* coming from quality approach.

3.2 *MATLAB*-BASED ALGORITHM FOR ADAPTIVE USABILITY EVALUATION OF E-LEARNING SYSTEMS

In adaptive usability evaluation of eLearning systems we propose in this study to calculate two different indexes as shown in figure 3.1.

Figure 3.1 UseLearn Model

One of them is calculated by following the first arrow. This index is named as *usability index* and is calculated by conventional usability dimensions which are mentioed in (Frokjaer, Hertzum & Hornbaek, 2000). They claim that *effectiveness* is defined as the accuracy and completeness with which users achieve certain goals. Indicators of effectiveness include quality of solution and error rates. *Efficiency* is the relation a) between the accuracy and completeness with which users achieve users certain goals and b) the resources expended in achieving them. Indicators of efficiency include task completion time and learning time. *Satisfaction* is the users' comfort and positive attitudes towards the use of the system. Users' satisfaction can be measured by attitude rating scales. In this study there were two tasks to perform using our eLearning system. Depending on these two tasks our modified checklist asked two questions (Q1 and Q2) to measure efficiency, two questions to measure effectiveness (Q3 and Q4). In fact, these are the questions asking for *the learning time*, *task completion time, the task completion percentage* and *error rates* respectively. The checklist asked the participants if they are satisfied with the eLearning course they received at the end of the checklist by taking the previous questions into account. This was the last question of checklist, Q45. This question aimed to measure the *satisfaction*.

Second index was calculated by following the second arrow. This index is named as *integrated index* because while calculating this index we took both eLearning usability and eLearning quality dimensions into account as mentioned in the previous part *2.1 Development of eLearning Usability Checklist*. Namely this index was an integrated index of usability and quality dimensions. In the figure above it is indicated how many questions were asked for each dimension. For instance the checklist asked three questions to measure the dimension *visibility*. These questions were Q5, Q6, and Q7. The other questions asked for the other dimensions can also be seen in the figure.

In both of these indexes the procedure which was followed was the same. To begin with, we can name the questions as subdimensions and we can assume that they join together and constitute the dimensions. For example, Q1 and Q2 (first and second questions) are subdimensions and these subdimensions constitute the dimension *efficiency*. Here we applied principal component analysis in Matlab. Let's consider efficiency first. In the first step we calculated the weights of Q1 and Q2, which were found as 0.47 and 0.53 respectively. Pay attention that for each separate dimension these weights must be 1.00 in total. Similarly, we found the weights of the other subdimensions also for integrated index dimensions. Second step was to find out the weights of the dimensions which constitute the indexes. For example, the dimensions of usability index are efficiency, effectiveness and satisfaction. These were calculated by principal component analysis and found as 0.20; 0.34 and 0.46 respectively. Similarly, the subdimension weights and the dimensions weights were calculated.

The followers may think what this method will work for. The answer of this question is explained in the figure 3.2. Adaptive usability evaluation is aiming at to find out how much the usability is and what must be done to increase this existing usability index. First of all, the evaluaters must gather data related to their eLearning system by the help of their checklist and the tasks in this checklist. After that, it is obvious that this data must be entered in any software program and analyzied. Here we propose to perform PCA in Matlab to find out the weights of the subdimensions and dimensions. Depending on these weigths, we can calculate the matrix of each latent dimension. By integrating these dimensions with neural networks, the indexes can be calculated. If the index scores are satisfactory, there is no need to go further. Of course, being satisfactory is a subjective decision. This can mean an increase comparing to the existing index score or there can be a threshold value. After this step, if the index score is satisfactory, the adaptive usability valuation process is stopped. If not, in the opposite direction of index calculation, firstly the most critical dimensions are searched for.

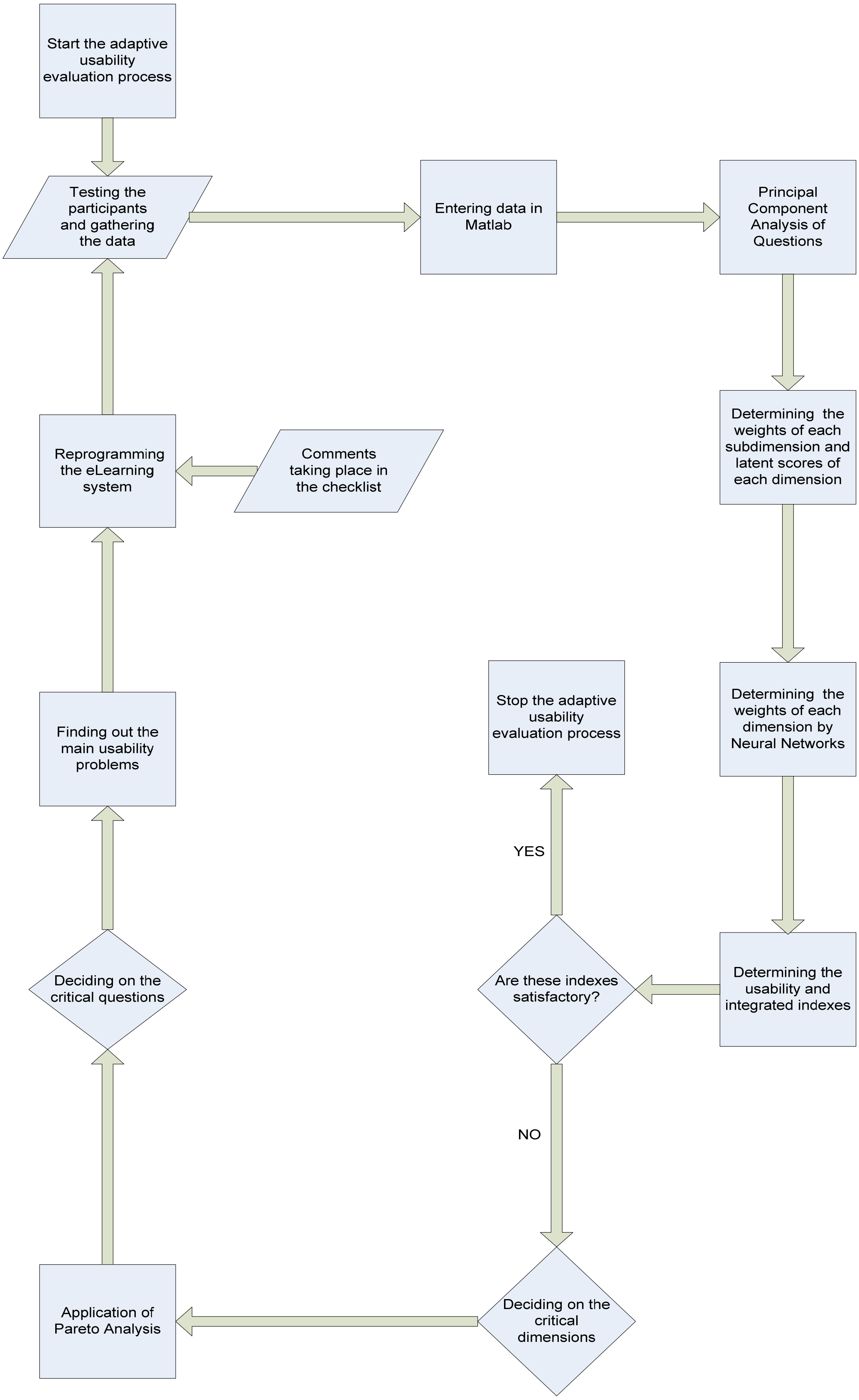

Figure 3.2 UseLearn Algorithm: Adaptive Usability Evaluation of eLearning Systems

Of course here again, the most critical dimensions are subjective and they can be determined by a threshold value or arbitrarily. After this decision, similarly the most critical subdimensions (questions) are determined. Depending on the problems that the questions indicate and the comments that were made by the participants, the improvements must be done and the eLearning system must be reprogrammed. Following this step, the eLearning system must be tested by the participants. Here it would be better to make the test to the same participants as the increase in the system can be noticed by them significantly.

It is quite obvious adaptive usability evaluation process is a never ending process to reach to the most usable eLearning system. But it is also too much time consuming and costly. So after some evaluation processes, it has to be stopped.

3.3 QUANTITATIVE EVALUATION TECHNIQUES OF USELEARN METHOD

Here we also would like to mention about the quantitative techniques that we used as our evaluation tool. These techniques are *Principal Component Analysis, Neural Networks* and *Pareto Analysis*.

3.3.1 Principal Component Analysis

One of the difficulties inherent in multivariate statistics is the problem of visualizing multidimensionality. When there are more than three variables, it stretches the imagination to visualize their relationships. Fortunately, in data sets with many variables, groups of variables often move together. One reason for this is that more than one variable may be measuring the same driving principle governing the behavior of the system. In many systems there are only a few such driving forces. But an abundance of instrumentation enables you to measure dozens of system variables. When this happens, you can take advantage of this redundancy of information. You can simplify the problem by replacing a group of variables with a single new variable. Principal components analysis is a quantitatively rigorous method for achieving this simplification. The method generates a new set of variables, called principal components. Each principal component is a linear combination of the original variables. All the principal components are orthogonal to each other so there is no redundant information. The principal components as a whole form an orthogonal basis for the space of the data. There are an infinite number of ways to construct an orthogonal basis for several columns of data. What is so special about the principal component basis? The first principal component is a single axis in space. When you project each observation on that axis, the resulting values form a new variable. And the variance of this variable is the maximum among all possible choices of the first axis. The second principal component is another axis in space, perpendicular to the first. Projecting the observations on this axis generates another new variable. The variance of this variable is the maximum among all possible choices of this second axis. The full set of principal components is as large as the original set of variables. But it is commonplace for the sum of the variances of the first few principal components to exceed 80% of the total variance of the original data. By examining plots of these few new variables, researchers often develop a deeper understanding of the driving forces that generated the original data Principal-component analysis(PCA) is a useful technique you can use to reduce the dimensionality of large data sets, such as those from microarray analysis. PCA can also be used to find signals in noisy data. In some situations, the dimension of the input vector is large, but the components of the vectors are highly correlated (redundant). It is useful in this situation to reduce the dimension of the input vectors. An effective procedure for performing this operation is principal component analysis. This technique has three effects: it orthogonalizes the components of the input vectors (so that they are uncorrelated with each other); it orders the resulting orthogonal components (principal components) so that those with the largest variation come first; and it eliminates those components that contribute the least to the variation in the data set. Note that we first normalize the input vectors, so that they have zero mean and unity variance. This is a standard procedure when using principal components. This means that the principal components that contribute to the total variation in the data set less than a determined percentage are eliminated. The matrix ptrans contains the transformed input vectors. After the network has been trained, this matrix should be used to transform any future inputs that are applied to the network. It effectively becomes a part of the network, just like the network weights and biases (Jolliffe, 1986).

3.3.2 Neural Networks

In this chapter we will partially mention about Neural Networks, about the parts that we used in our case study analysis.

- *Feed-Forward Backpropagation* Network Type:

Backpropagation can train multilayer feed-forward networks with differentiable transfer functions to perform function approximation, pattern association, and pattern classification. (Other types of networks can be trained as well, although the multilayer network is most commonly used.) The term backpropagation refers to the process by which derivatives of network error, with respect to network weights and biases, can be computed. This process can be used with a number of different optimization strategies.

newff creates a feed-forward backpropagation network. Feed-forward networks consist of Nl layers using the dotprod weight function, netsum net input function, and the specified transfer functions. The first layer has weights coming from the input. Each subsequent layer has a weight coming from the previous layer. All layers have biases. The last layer is the network output. Each layer's weights and biases are initialized with initnw. Adaption is done with trains, which updates weights with the specified learning function. Training is done with the specified training function. Performance is measured according to the specified performance function.

As noted in Neuron Model and Network Architectures, for multiple-layer networks we use the number of the layers to determine the superscript on the weight matrices. The appropriate notation is used in the two-layer `tansig/purelin` network shown in figure 3.3.

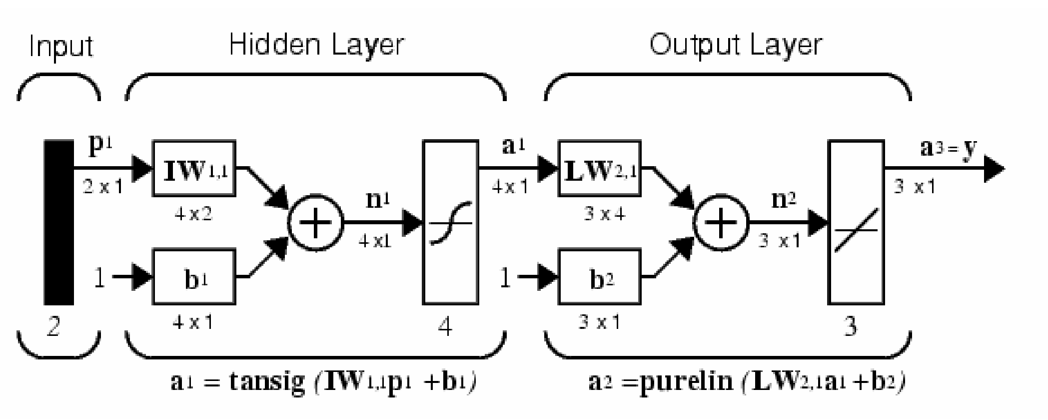

Figure 3.3 Two-Layer *Tansig/Purelin* Neural Network (adapted from Matlab 7.0 tutorial)

This network can be used as a general function approximator. It can approximate any function with a finite number of discontinuities, arbitrarily well, given sufficient neurons in the hidden layer.

- *tansig* Transfer Function:

tansig is a transfer function. Transfer functions calculate a layer's output from its net input. tansig(N) takes one input,

N -- S x Q matrix of net input (column) vectors

and returns each element of N squashed between -1 and 1. tansig(code) return useful information for each code string:

'deriv' -- Name of derivative function

'name' -- Full name

'output' -- Output range

'active' -- Active input range

tansig is named after the hyperbolic tangent, which has the same shape. However, tanh may be more accurate and is recommended for applications that require the hyperbolic tangent.This is a hyperbolic tangent sigmoid transfer function shown in figure 3.4.

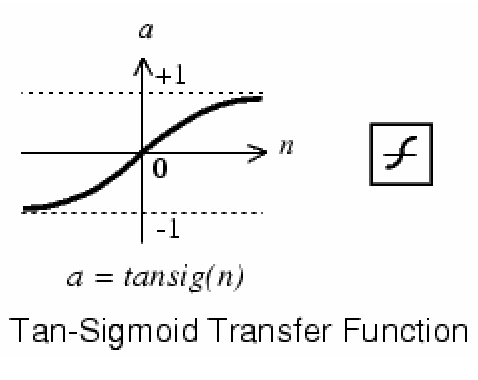

Figure 3.4 Tan-Sigmoid Transfer Function (adapted from Matlab 7.0 tutorial)

tansig(N) calculates its output according to: n = 2/(1+exp(-2*n))-1

This is mathematically equivalent to tanh(N). It differs in that it runs faster than the MATLAB implementation of tanh, but the results can have very small numerical differences. This function is a good trade off for neural networks, where speed is important and the exact shape of the transfer function is not.

- *purelin* Transfer Function:

purelin is a transfer function which can be represented as in the figure 3.5. Transfer functions calculate a layer's output from its net input. purelin(N) takes one input,

N -- S x Q matrix of net input (column) vectors

and returns N.

purelin(code) returns useful information for each code string:

'deriv' -- Name of derivative function

'name' -- Full name

'output' – Output range

'active' – Active input range

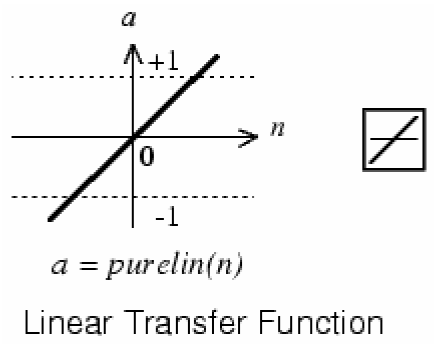

Figure 3.5 Linear Transfer Function (adapted from Matlab 7.0 tutorial)

You can create a standard network that uses purelin by calling newlin or newlind. To change a network so a layer uses purelin, set net.layers{i}.transferFcn to 'purelin'. In either case, call sim to simulate the network with purelin.

- *trainlm* Training Function:

trainlm is a network training function that updates weight and bias values according to Levenberg-Marquardt optimization.

trainlm can train any network as long as its weight, net input, and transfer functions have derivative functions. Backpropagation is used to calculate the Jacobian jX of performance perf with respect to the weight and bias variables X. Each variable is adjusted according to Levenberg-Marquardt, jj = jX * jX

je = jX * E

dX = -(jj+I*mu) \ je

where E is all errors and I is the identity matrix. The adaptive value mu is increased by mu_inc until the change above results in a reduced performance value. The change is then made to the network and mu is decreased by mu_dec. The parameter mem_reduc indicates how to use memory and speed to calculate the Jacobian jX. If mem_reduc is 1, then trainlm runs the fastest, but can require a lot of memory. Increasing mem_reduc to 2 cuts some of the memory required by a factor of two, but slows trainlm somewhat. Higher values continue to decrease the amount of memory needed and increase training times. Training stops when any of these conditions occur: The maximum number of epochs (repetitions) is reached. The maximum amount of time has been exceeded. Performance has been minimized to the goal. The performance gradient falls below mingrad. mu exceeds mu_max. Validation performance has increased more than max_fail times since the last time it decreased when using validation (Vogl et al., 1988).

3.3.3 Pareto Analysis

A defect or nonconformity is the unit of a product that does not satisfy one or more specifications for a product; that is, each specific point at which a unit does not meet the specification is known as a defect or nonconformity (Montgomery, 1997). This type of data is informative because there will usually be more than one type of nonconformity. By analyzing the defects by type, you often gain valuable information about the cause.

In Montgomery's printed circuit board example, there were 16 different types of defects. Plotting those defects on a Pareto Chart points out the major types of the defects (those that occur most frequently).

The Pareto Analysis attempts to sort out the vital few causes to gain an understanding of where to begin an improvement process. The analysis plots the frequency of each defect type, which helps identify the most frequently occurring type of defects. The analysis does not identify the most important defects; it simply reveals the most frequent defects.

The Pareto Chart is a histogram for categorical data. This chart, together with attributes control charts, helps to identify and rank order the most important causes of a problem within a process, or to track the progress of changes to a process (Montgomery, 1997).

3.4 SUMMARY OF RESULTS OF CHAPTER 3

In chapter 3, a checklist including both quality and usability dimensions was developed. Based on the checklist, a common quantitative adaptive usability evaluation method *UseLearn* was proposed. It includes many different quantitative techniques such as principal components analysis, neural networks and pareto analysis.

CHAPTER 4

CASE STUDY

4.1 DESCRIPTION OF AN E-LEARNING SYSTEM IN CELL BIOLOGY

In this study, an eLearning biology course was examined. Cell biology course was selected as the eLearning course tool. Because it would be better if we applied a qualitative course not a quantitative one. A qualitative course would be easier to understand by the help of an eLearning system rather than a quantitative one. We also decided to use the graphics and figures broadly to make use of eLearning system. The most suitable course fitting to these requirements was biology.

For presenting this biology course, an open area eLearning system "*moodle*" was used as the eLearning tool.

Moodle is a VLE developed and written by a PhD student called Martin Dougiamas who was interested in creating a "social constructionist framework" of education within a computerised system.

Platform: Apache, PHP, MySQL/PostgreSQL.

Installing Moodle was relatively simple. It required the source files to be downloaded from the Moodle Website, and then decompressed onto the local hard disk. Once it was successfully saved, the required files were transferred to the web server and then the settings were changed in the configuration files to match our settings at Progress through Training. The installation of Moodle was aided by several automated pages, which speeded up the installation process. (Moodle, 23.6.2005)

Installation guidelines (taken from the installation package):

Move the Moodle files into your web directory.

- Create an empty directory somewhere to store uploaded files (NOT accessible via the web).

This directory must be writeable by the web server process.

- Create a single database for Moodle to store all it's tables in (or choose an existing database).
- Copy the file config-dist.php to config.php, and edit it with all your own settings.
- Visit your new home page with a web browser. Moodle will lead you through the rest of the Set-up, creating an admin account and so on.

- Set up a cron task to call the file admin/cron.php every five minutes or so.

4.2 USABILITY EVALUATION OF FIRST VERSION OF E-LEARNING SYSTEM IN CELL BIOLOGY

4.2.1 Planning and Preparing the Experiment

A user friendly interface was created for usability testing of eLearning system with the aid of the web module “moodle” created by PHP and HTML. MYSQL database software was used with the programs mentioned above to store any type of data such as; user information, courses, tests, test scores, timing information, etc as shown in figure 4.1.

Figure 4.1 Planning the experiment in Moodle, Screenshot 1

The graphical user interface was created and published on the internet, in the name of "eLearning Usability in Biology" as can be seen in figure 4.2.

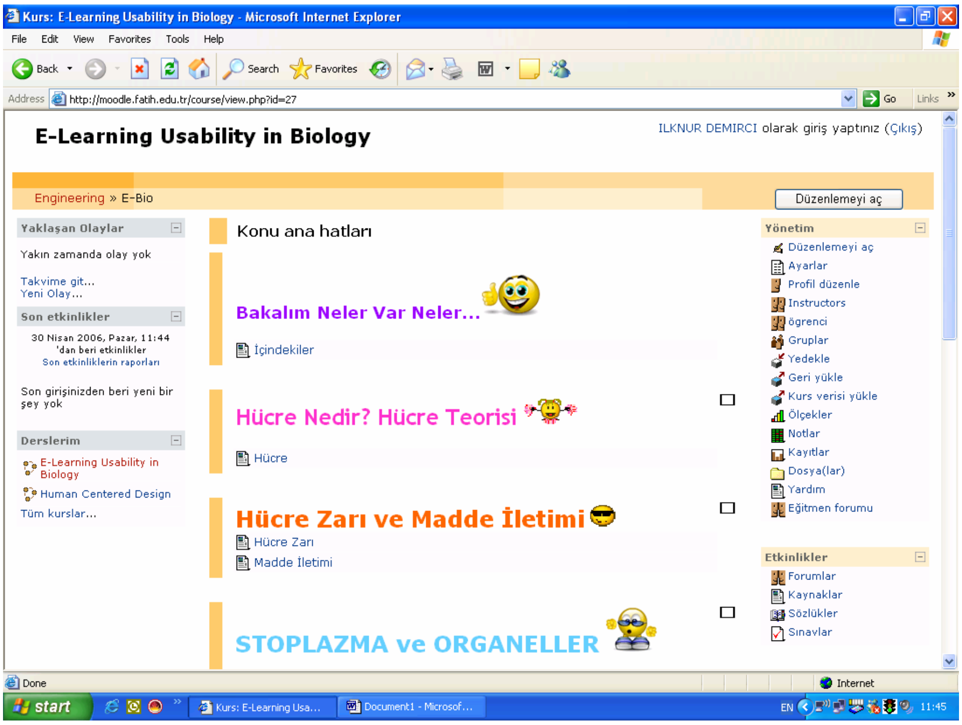

Figure 4.2 Planning the experiment in Moodle, Screenshot 2

The system contains four main sections; table of content, lectures, test and dictionary as it can be seen in figure 4.3.

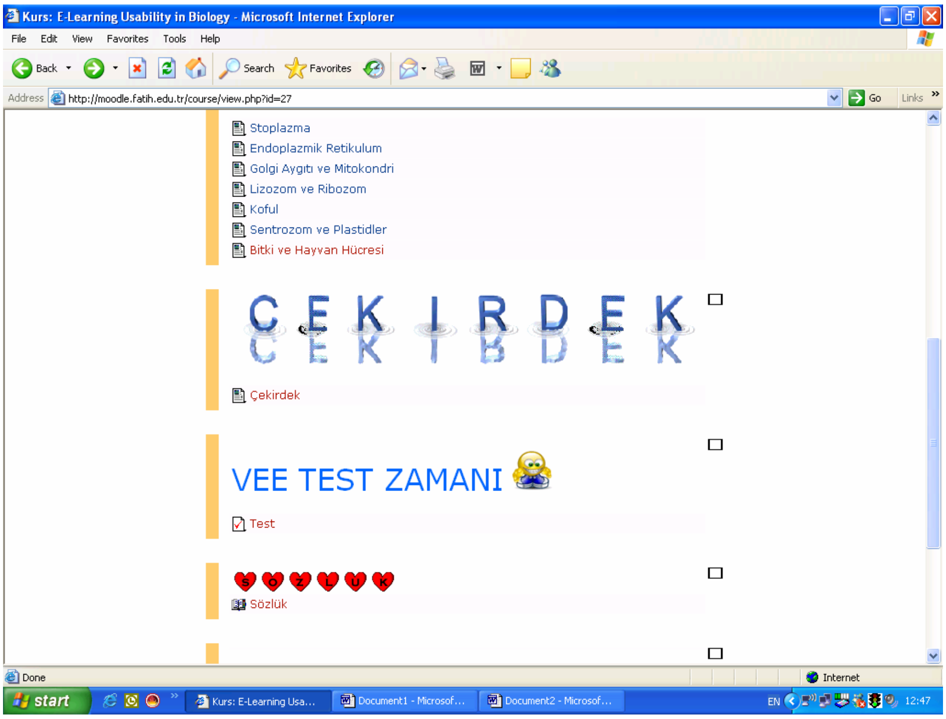

Figure 4.3 Planning the experiment in Moodle, Screenshot 3

Flash animated items were used to increase the user friendliness of the system. Each item of the content table is designed as links to enhance the accessibility of the lectures. Hence access to the lectures both from the main page as well as the table of content was provided. (See figure 4.4).

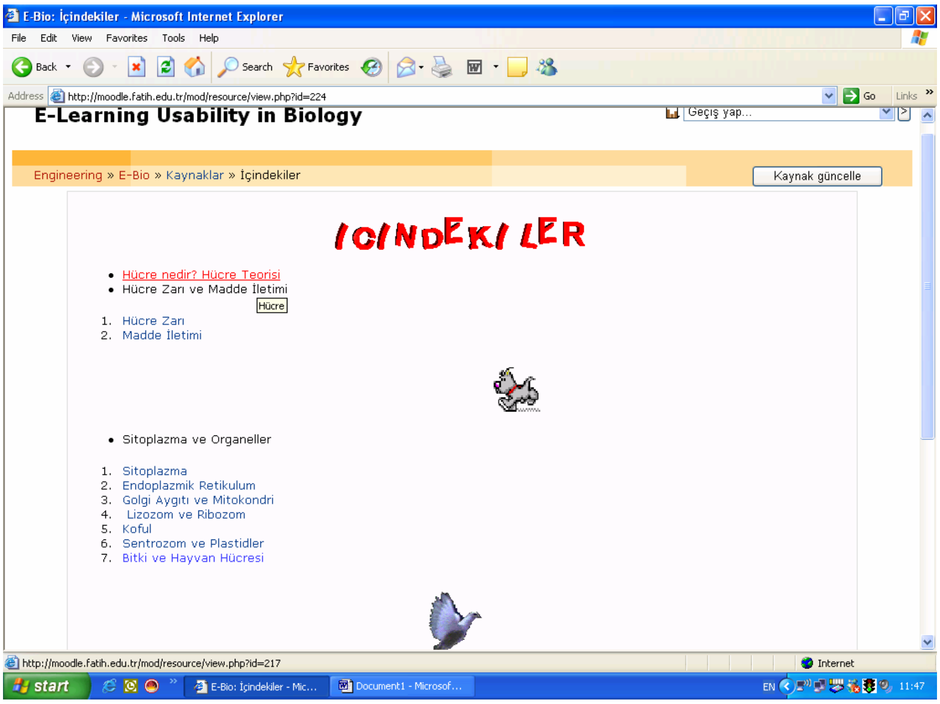

Figure 4.4 Planning the experiment in Moodle, Screenshot 4

Lectures were designed with descriptive pictures providing content enrichment and better understanding of the subject. Different colors and highlights were used to draw the student's attention and emphasize the important points. Most of the figures and pictures were designed as moving animations to make learnability more effective. (One of these animations was shown in figure 4.5).

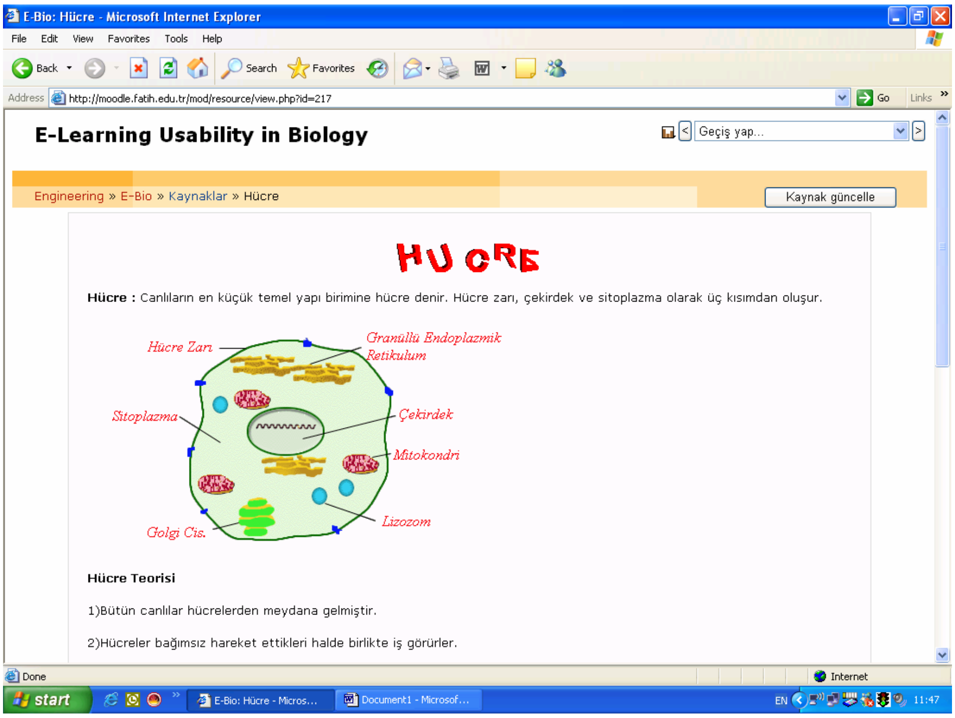

Figure 4.5 Planning the experiment in Moodle, Screenshot 5

A dictionary was created to ease the understanding of complex terms and prevent any ambiguities, which might be possible cause of confusions. This section was created dynamically as to be updated with new words to help further studies and clear understanding of the subject. A view of the dictionary in shown in figure 4.6.

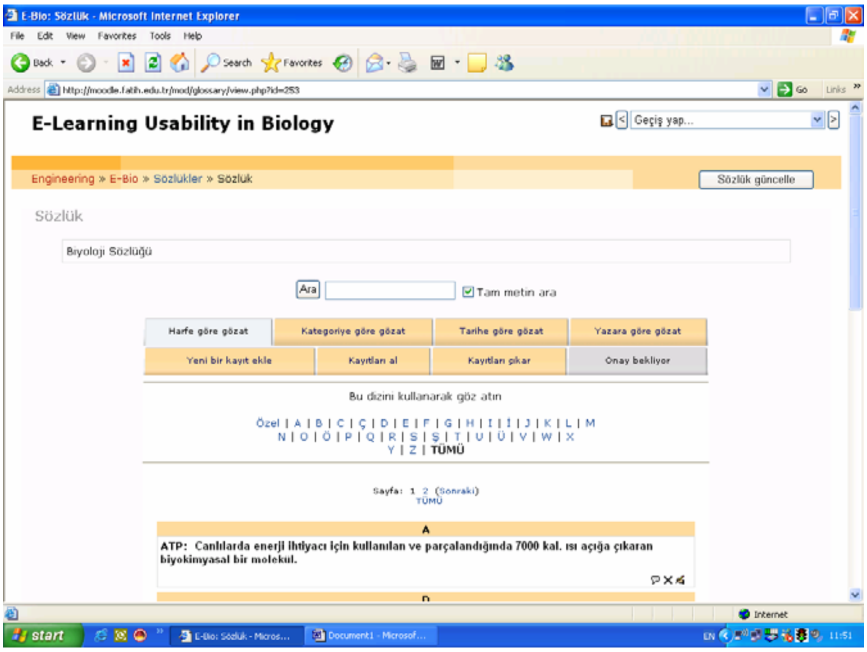

Figure 4.6 Planning the experiment in Moodle, Screenshot 6

After finishing these preparations, we determined whom to test. For gathering more reliable data, we requested the teachers to encourage the students somehow. The way that we followed was to promise some extra credits for their grading approximately (5 or 10 %). But this promise was valid only if they answered the questions in the statistical assurance intervals of 90 %. Depending on these prewarnings the results were reliable enough and we didn't need to delete any of them.

4.2.2 Description of Experiment

The experiment held in this study was two stepped. Adaptive usability evaluation process which is mentioned in part 2.2 was followed. First, checklist tool which is mentioned in part *2.1 Modified eLearning Usability Checklist* was experimentally studied by testing and evaluating the eLearning system by using a biology course in moodle. Secondly, depending on the data processing and the checklist comments of the participants, the major usability problems were found out. In the second step the improvements were carried out, the eLearning system reprogrammed and again data were collected. These two steps are explained in details in the following.

In our modified checklist apart from the questions related to eLearning usability and quality dimensions, some other questions were also asked to measure *efficiency*, *effectiveness*, and *satisfaction*. To measure efficiency, checklist asked *the learning time* and *task completion time* as mentioned in (Froakjaer, Hertzum & Hornbaek, 2000) and to measure the effectiveness, checklist asked *the task completion percentage* and *error rates*. There were two tasks that the test participants were expected to solve. Performing the required two tasks and answering the questions of the checklist took approximately 60-70 minutes per student. This was a relatively time consuming study but this was a naturel result of the requirement that the participants had to follow the biology course content and then have a quiz related to this content.

As mentioned in part *2.1 Development of eLearning usability checklist*, this was a modified checklist of usability and quality dimensions related to eLearning. Moreover, although participant's English knowledge was good enough to perform the tasks, to make the questions more understandable they were translated also in Turkish. Therefore, we applied a pretest to 5 students to find out some translation problems and confusing questions. Depending on their reactions, we deleted some questions and changed some of the translations if possible. Preceding these changes we applied the test to our real data set.

The set of the participants was a good mixture of two parts. These parts and their percentages were as following: 48 % high school students of Fatih College, 52 % university students of Fatih University. 84 % of the university students were in the Department of Biology and the rest were in the Department of Industrial Engineering. 64 % of all participants were male and 36 % of them were female. In the first step of the data collection, totally 88 people were examined.

All tests were carried out in the computer laboratories of Fatih University and Fatih College by allocating one computer to each participant. The number of the people working around and the noise level were nearly the same.

The participants received checklist documents containing the aim of the survey, how they would proceed and some descriptive questions such as their names, departments, if they are university or high school student and gender. They were asked to perform two tasks using the eLearning biology course in moodle which is supported by Fatih University Faculty of Engineering. These tasks were the same for all participants. After the session they were asked to fill out the checklist.

4.2.3 Results of UseLearn-based Analysis

The data gathered in the experiment was entered to Matlab first to run principal component analysis to find out the weights of the subdimensions. The names of the dimensions, their abbreviations, the weights of the subdimensions and the score of integrated index gotten by PCA are shown in the table 4.1 and in table 4.2. In the appendicies part of the thesis we also would like to present the Eigenvalue scores and percent of varience explained by each variable to prove the reliability of the principal component analysis that we performed. Please see appendicies section for these further information.

Table 4.1 The Names and the Abbreviations of the Dimensions

The name of the dimension	The abbreviation of the dimensions
Visibility	V
Aesthetics	AE
Course Management	CM
Interactivity	I
Consistency&Functionality	CF
Feedback&Help	FH
Memorability	M
Completeness	COM
Accessibility	AC
Flexibility	FL
Reducing Redundancy	RE
Error Prevention	EP
Clear Syllabus	CS
Variety of Learning Tasks	VLT
Assessment Strategy	AS

Table 4.2 The Weights of the Subdimensions of Integrated Index Version 1

The subdimensions of integrated index	The weights of each subdimension
V1 (Q5)	0.34
V2 (Q6)	0.32
V3 (Q7)	0.34
AE1 (Q8)	0.32
AE2 (Q9)	0.26
AE3 (Q10)	0.42
CM1 (Q11)	0.53
CM2 (Q12)	0.47
I (Q13)	1
CF1 (Q14)	0.30
CF2 (Q15)	0.23
CF3 (Q16)	0.20
CF4 (Q17)	0.27
FH1 (Q18)	0.28
FH2 (Q19)	0.20
FH3 (Q20)	0.52
M1 (Q21)	0.36
M2 (Q22)	0.26
M3 (Q23)	0.19
M4 (Q24)	0.18
COM1(Q25)	0.60
COM2 (Q26)	0.40
AC1 (Q27)	0.12
AC2 (Q28)	0.47
AC3 (Q29)	0.42
FL1 (Q30)	0.89
FL2 (Q31)	0.11
RE1 (Q32)	0.49
RE2 (Q33)	0.23
RE3 (Q34)	0.28
EP1 (Q35)	0.27
EP2 (Q36)	0.34
EP3 (Q37)	0.39
CS1 (Q38)	0.35
CS2 (Q39)	0.35
CS3 (Q40)	0.30
VLT1 (Q41)	0.52
VLT2 (Q42)	0.48
AS1 (Q43)	0.43
AS2 (Q44)	0.57
Integrated Index Score	0.69

In table 4.2, note that the total weights of subdimensions of a spesific dimension is equal to 1. For example FH1 (0.28) + FH2 (0.20) + FH3 (0.52) =1. There is no subdimension for *interactivity*, in other words the dimension *interactivity* is measured by only one question (subdimension) in the checklist so its weight is directly equal to 1.

Similarly, we calculated the weights of dimensions and subdimensions of usability index and the total usability index score in PCA. The results are shown in the table 4.3 and in table 4.4.

Table 4.3 The Weights of the Subdimensions of Usability Index Version 1

The subdimensions of Usability Index	The weights of the subdimensions
Efficiency1 (Q1)	0.47
Efficiency2 (Q2)	0.53
Effectiveness1 (Q3)	0.61
Effectiveness2 (Q4)	0.39
Satisfaction (Q45)	1

Table 4.4 The Weights of the Dimensions of Usability Index Version 1

The dimensions of Usability Index	The weights of the dimensions
Efficiency	0.20
Effectiveness	0.34
Satisfaction	0.46
Usability Index Score	0.59

Depending on PCA latent scores for each dimension we ran Neural Networks in MATLAB starting with 3 neurons ending at 14. The inputs were 15 latent scores of dimensions of integrated index and the output was the usability index, all of which were calculated by PCA. There were two layers. The number range of neurons in layer one was determined depending on the number of dimensions we have. Here there were 15 dimensions so we stopped at the number of 14 neurons. In the second layer one neuron was settled because we had only one output. The summary of the NN model of Uselearn approach is shown in table 4.5.

Table 4.5 NN Model of UseLearn Approach for the Case Study

Network Type:	Feed-forward Backpropatation
Training Function:	Trainlm
Performance Function:	Mean Squared Error (MSE)
Transfer Function:	Layer 1-- Tansig
	Layer 2--Purelin

The same neural network models shown in table 4.5 were run in Matlab with only one difference: the number of neurons in Layer 2 they include; starting with 3 neurons in Layer 2 ending up with 14 neurons. The summary of these neural netwok models are shown in table 4.6.

Table 4.6 The Summary of NN Models in UseLearn Case Study Version 1

Number of Neurons	MSE of Testing Data	MSE of Training Data	Epoch
3	2.3444	0.0088	100
4	8.32E–01	7.45E–05	100
5	5.1893	0.0042	100
6	3.83E–01	3.48E–23	47
7	8.75E–01	1.11E–26	94
8	2.13E–01	7.52E–04	100
9	5.45E–01	6.78E–29	40
10	2.83E–01	1.33E–30	16
11	2.46E–01	4.82E–25	36
12	1.81E–01	1.66E–29	38
13	1.27E–01	2.04E–23	28
14	1.74E–01	1.37E–23	25

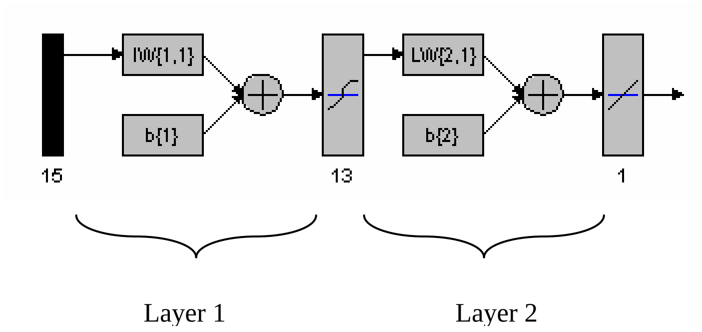

Figure 4.7 The View of the NN UseLearn Model for the Case Study Version 1

As seen in table 4.6 minimum captured MSE was performed with 13 neurons. So we can claim that our NN-Uselearn model is explained best with the model as in figure 4.7. The model parameters of this NN-Uselearn model with 13 neurons are given in figure 4.8.

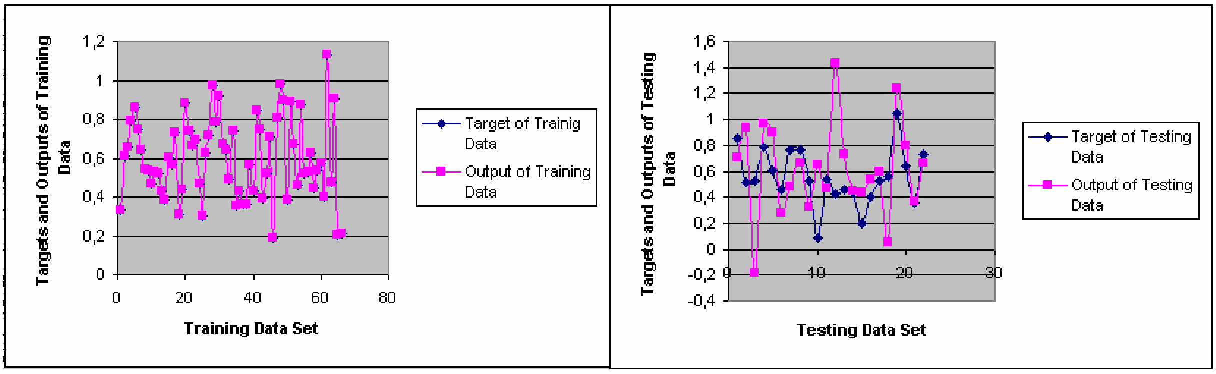

Figure 4.8 Targets and Outputs of Training and Testing Data Set Version 1

As the figure 4.8 indicates, we randomly assigned one fourth of the data (22 observations) as testing data and three fourth of it (66 observations) as training data set.

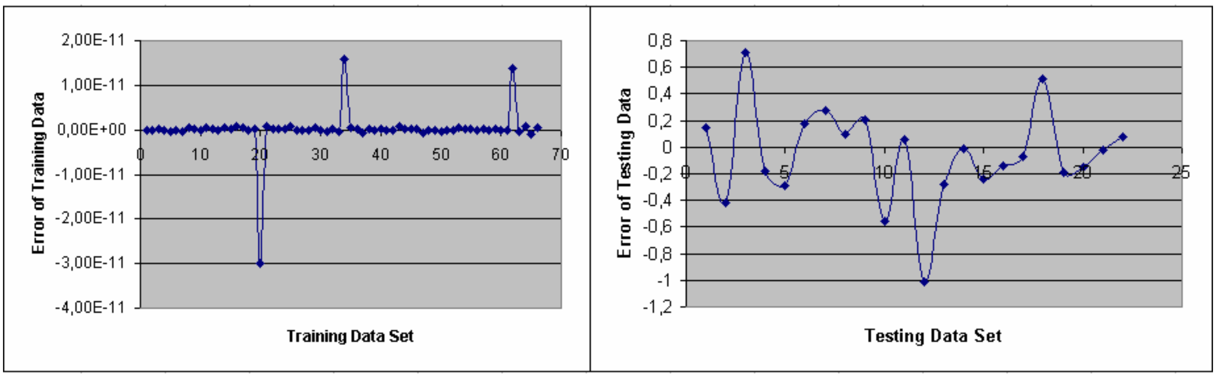

Figure 4.9 Error of Training and Testing Data Set Version 1

Error rates of the training data and testing data was found satisfactory. Thre were only a few outliers in training data and testing data errors were fluctuating around zero as shown in figure 4.9

Another screenshot of Matlab represents the changing graph of *trainlm* training function as shown in figure 4.10 with the error received and the epoch reached to receive that error.

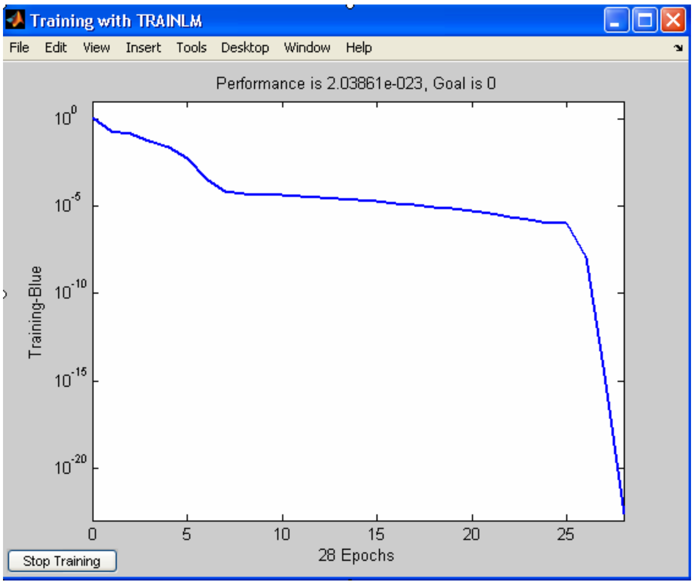

Figure 4.10 A View of Performance and Number of Epoch of *trainlm* training Function Version 1

The weights of the dimensions found with this model are shown in the table 4.7. We had to find out the most critical dimensions using these weights. But what would the threshold value be to stop tackling with the critical dimensions? To answer this question we performed Pareto Analysis.

Table 4.7 The Weights of the Dimensions of Integrated Index Version 1

The abbreviation of the dimensions	The weights of the dimensions by NN
V	0.170
AE	0.046
CM	0.049
I	0.040
CF	0.021
FH	0.001
M	0.125
COM	0.018
AC	0.038
FL	0.011
RE	0.005
EP	0.261
CS	0.178
VLT	0.028
AS	0.009

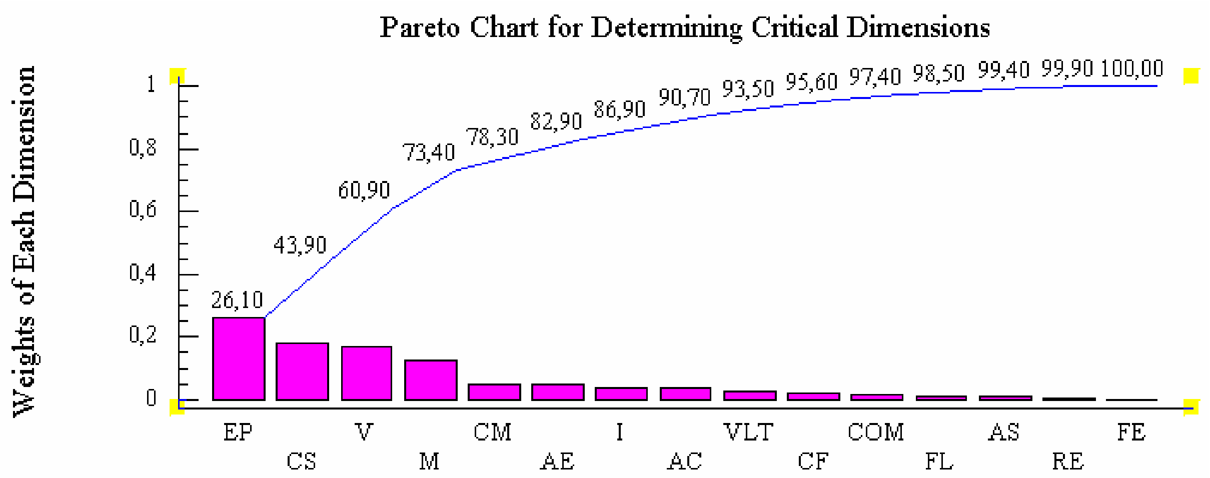

Figure 4.11 Pareto Analysis for Determining Critical Dimensions Version 1

In Pareto analysis, it is assumed that 80 % of the costs are caused by 20 % of the problems. Similarly, we performed here to find out the most critical dimensions affecting on eLearning usability of cell biology. As seen in figure 4.11, in our model the most critical 5 dimensions (which are 5/15=0.33, 33 % of all dimensions) affect the system 78.30 %. These dimensions and their weights in parenthesis were *Error Prevention (0.261), Clear*

Syllabus (0.178) Visibility (0.170), Memorability (0.125) and Course Management (0.049). So we decided to improve these dimensions to gain a higher index score.

4.2.4 Usability Problems and Design Recommendations

Next step was to find out to handle with which subdimensions (questions) of these five dimensions. For each dimension, we focused on the subdimensions whose sum of weights (or only weight) made up 20 % or more just to go on our hypothesis related to Pareto Analysis defining the threshold value. Taking the table ???? into account these were decided to be as following. Take care of that the weights in parenthesis are 20 % or more:

EP3: Do error or warning messages prevent possible errors from occuring? (0.39)
CS1: Are learning outcomes and performance expectations clearly defined? (0.35)
V1: Are options (buttons and selections) logically grouped and labeled? (0.34)
V3: Is course content meaningfully arranged with links from the homepage? (0.34)
M1: Is the user offered sufficient FAQ system and human support to obtain necessary help?(0.36)
CM1: Does the course contain important for the online students and link to the support areas?(0.53)

These problems were the ones which we found out by NN-Uselearn algoritm. But in our adaptive usability evaluation process we propose to take care of the participant comments in the checklist to improve and reprogram the eLearning system. Including these comments and usability problems we did the following improvements and reprogrammed our eLearning system.

Animations in the headers were removed; color and size of the fonts were changed to give the pages a common look.

- Analyzing the questionnaires it was found that animations were drifting the students' attention away from the content of the pages. Therefore the headers and some of the pictures in the lectures were made into a simpler form without animations. The previous design is shown in figure 4.12 and the second improved version is seen in figure 4.13.

- Font colors and sizes were changed into a common form within each section as well as in general.

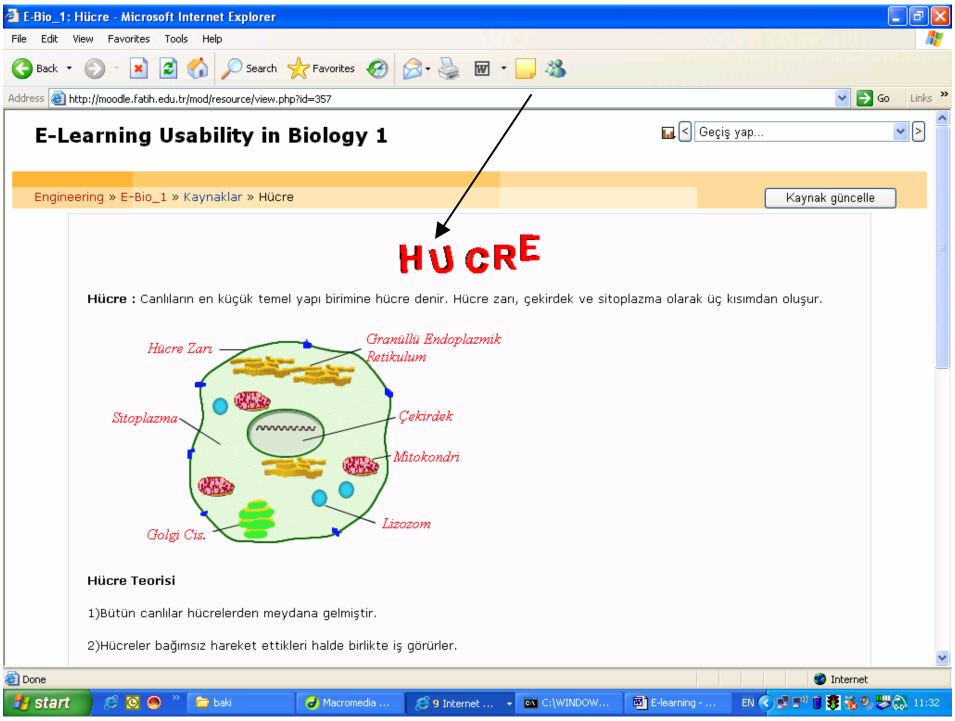

Figure 4.12 Improvements for Second Version, Screenshot 1

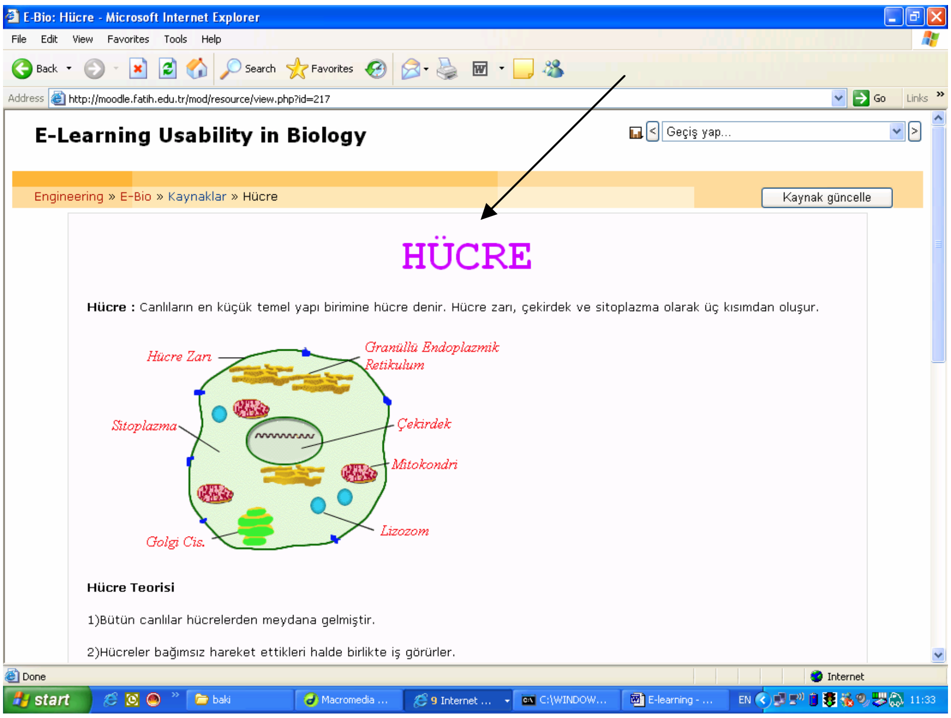

Figure 4.13 Improvements for Second Version, Screenshot 2

• Previously, navigation between the pages was done through a drop down list or from the index page only as shown in figure 4.14 but this was not convenient for many users.

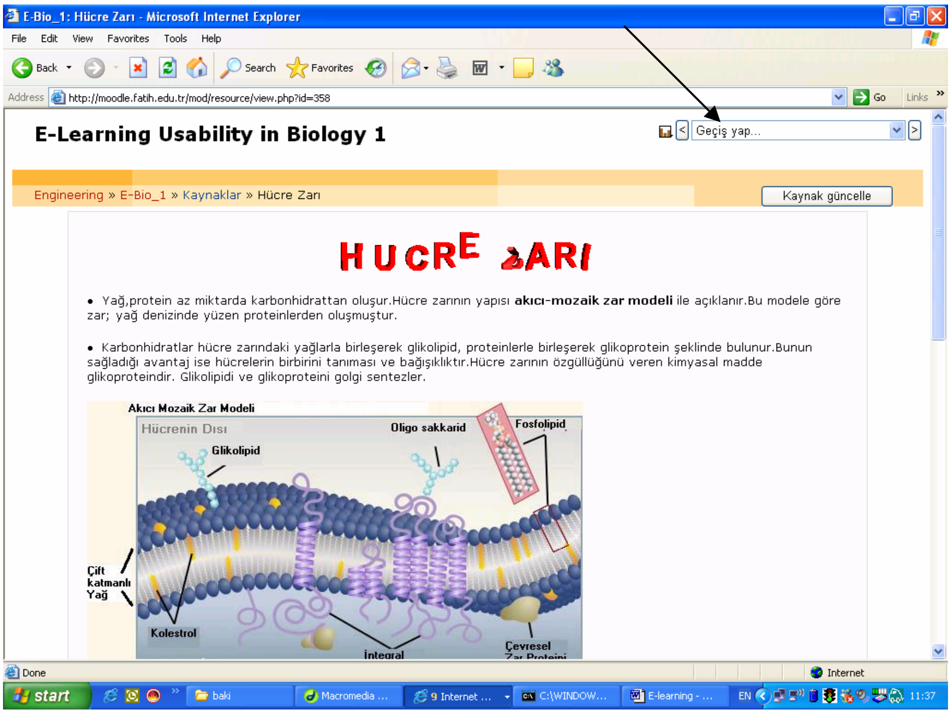

Figure 4.14 Improvements for Second Version, Screenshot 3

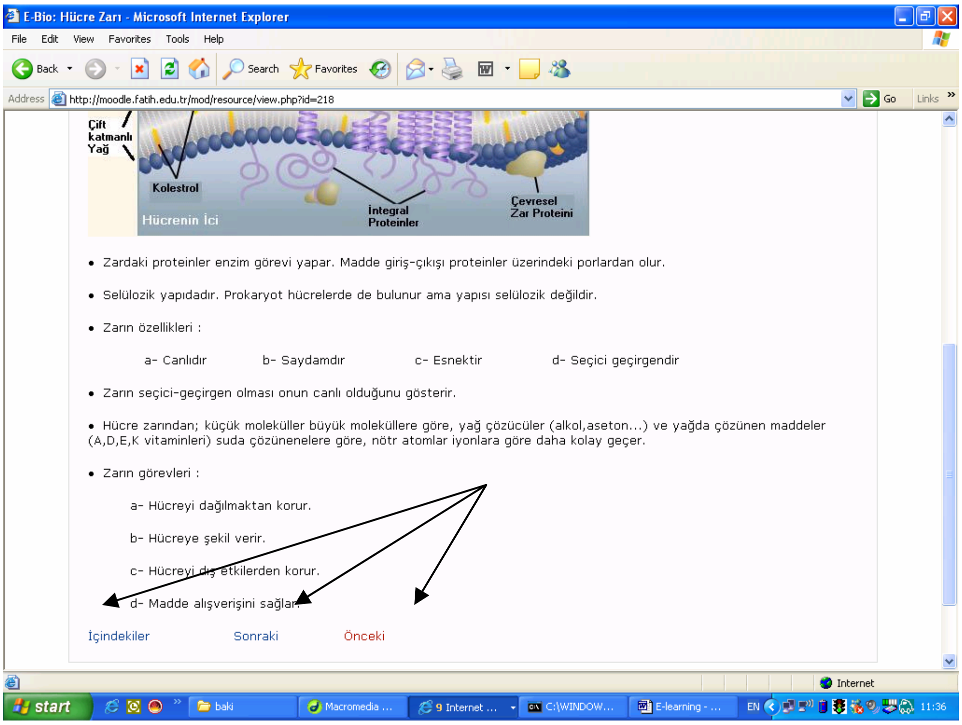

Figure 4.15 Improvements for Second Version, Screenshot 4

In order to ease the navigation, previous, back and index links were added to the bottom of every lecture page as it can be seen in the figure 4.15.

Numberings, bulletins and fonts of the choice lists in the test were changed into a uniform type for all the questions.

- The pre-choices in the questions were designed with different fonts and numberings previously, all of them were changed to be in the same font and numbering/bulletin style in the updated version.
- Unnecessary spacing between the questions was eliminated in order to avoid long scroll downs as shown in figure 4.16.

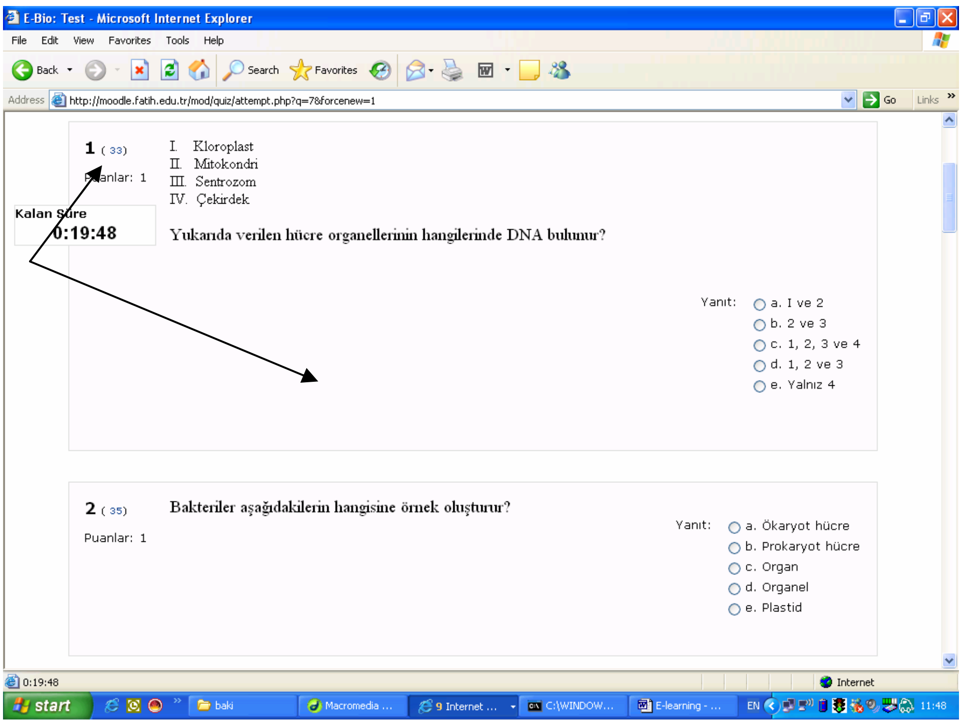

Figure 4.16 Improvements for Second Version, Screenshot 5

Previously underlined words were changed into a greater size with bold characters maintaining the emphasize on the questions and make the readability easier as shown in figure 4.17.

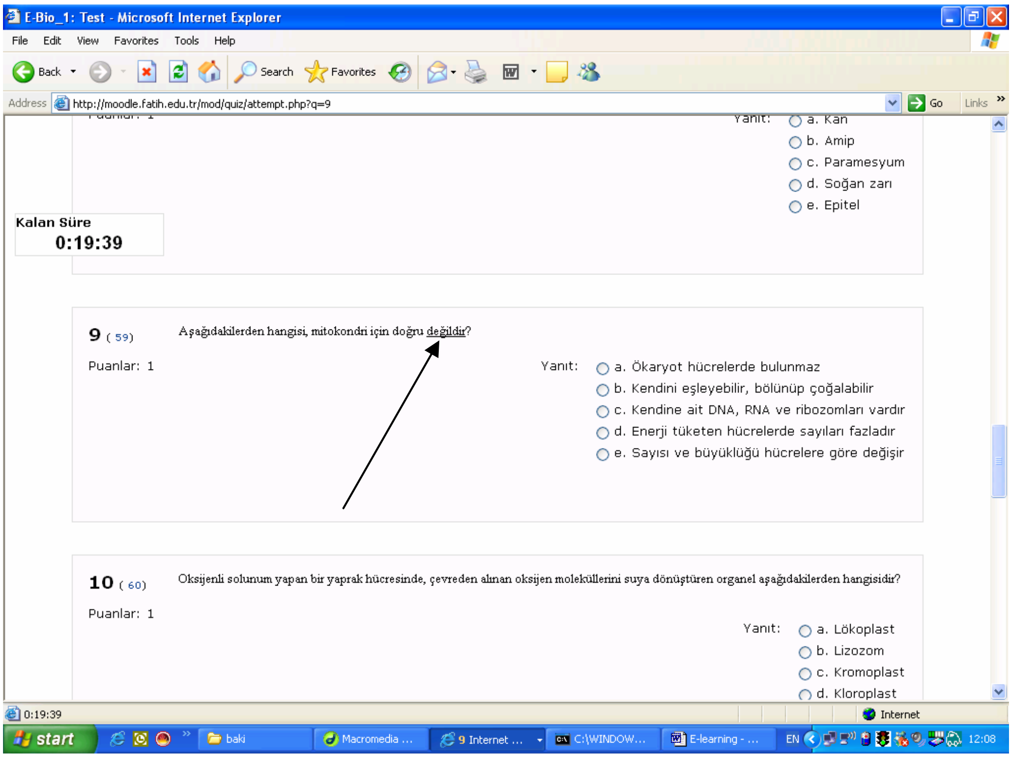

Figure 4.17 Improvements for Second Version, Screenshot 6

In order to have greater emphasize with notice some words were underlined in the questions. Since it is not allowed to have underlined words in a web page, they were removed and the emphasize was maintained by using bigger font size and bold characters which can be seen in figure 4.18.

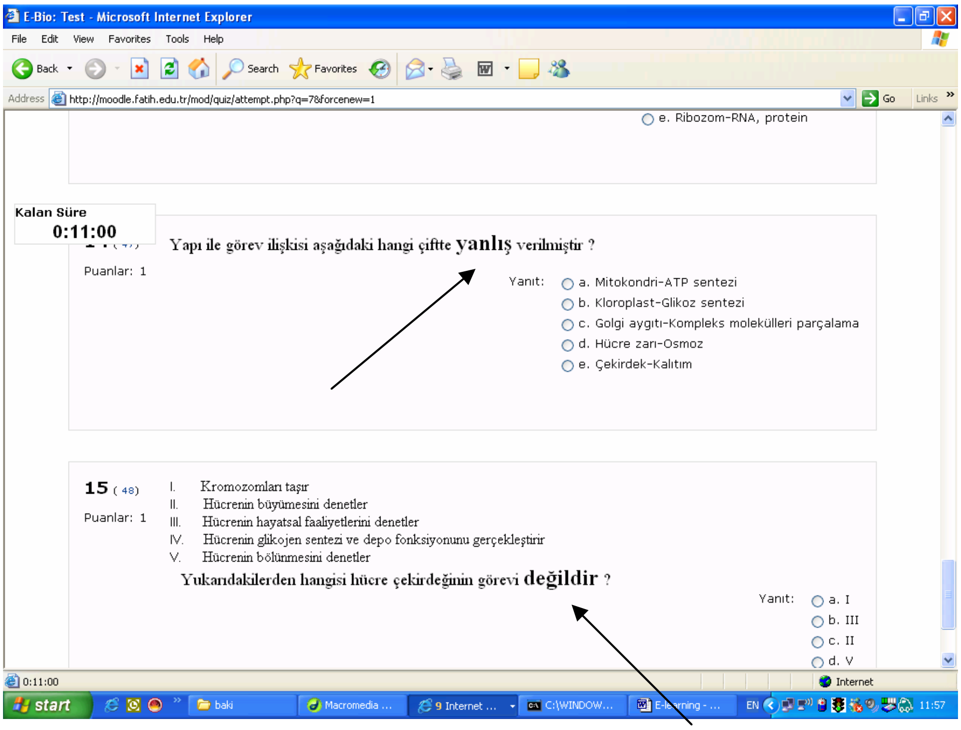

Figure 4.18 Improvements for Second Version, Screenshot 7

Useful links for e-learning and a section for FAQ were added. Note that FAQ is SSS in Turkish abbreviation. Both of these improvements are available in figure 4.19.

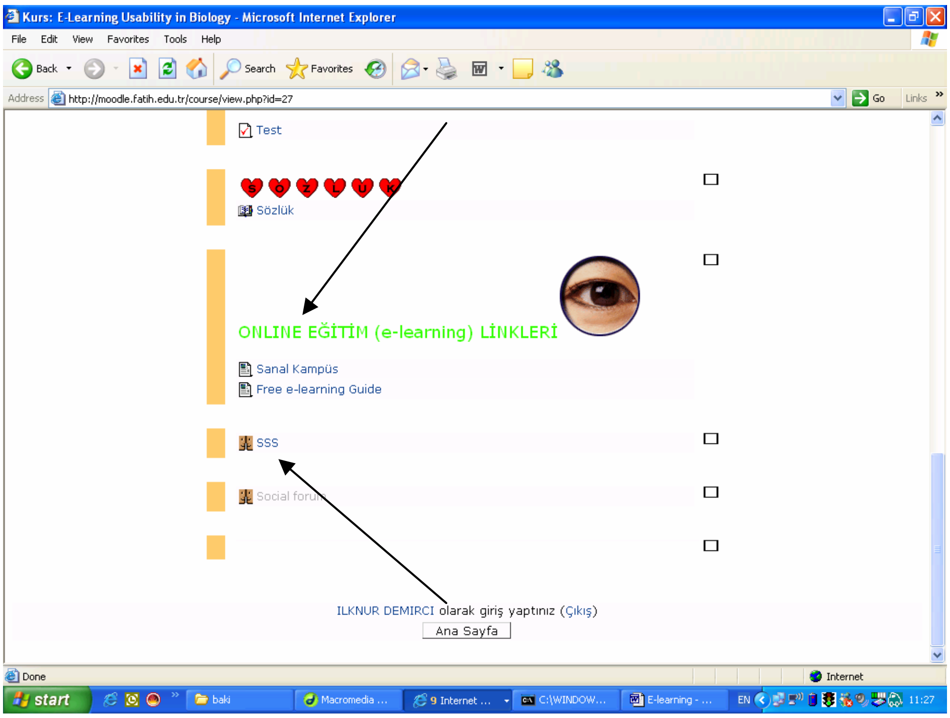

Figure 4.19 Improvements for Second Version, Screenshot 8

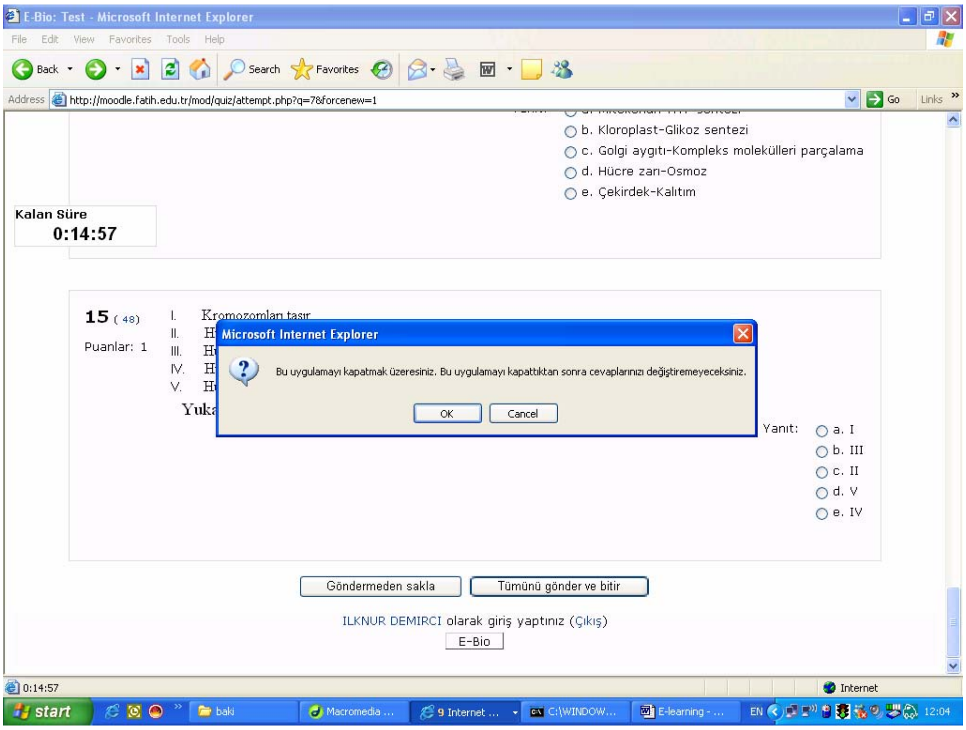

Figure 4.20 Improvements for Second Version, Screenshot 9

Warning messages were added to improve the usability as shown in figure 4.20.

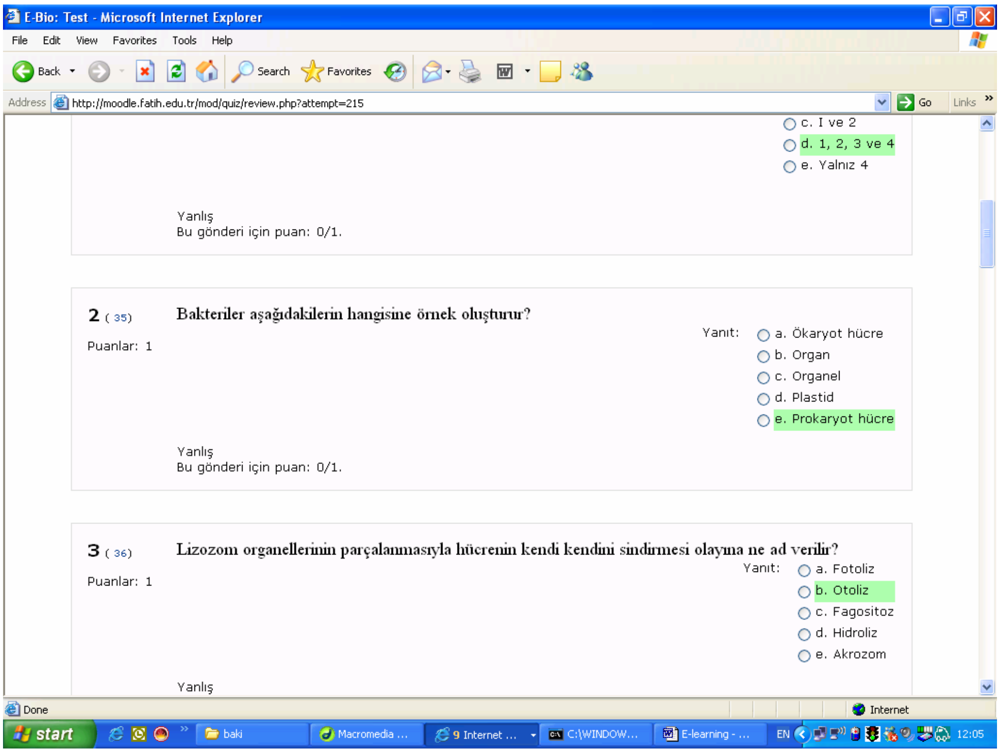

Figure 4.21 Improvements for Second Version, Screenshot 10

The correct answers were highlighted in green color in order to provide a comparison for the right and wrong answers of the test that is held at the end of the online course. This improvement is also shown in figure 4.21.

4.3 USABILITY EVALUATION OF SECOND IMPROVED VERSION OF E-LEARNING SYSTEM IN CELL BIOLOGY

4.3.1 Results of UseLearn-based Analysis

After performing the improvements in the online cell biology course, we retested the participants and this time the dataset size was 72 students, all from Fatih College. 60% were male and the rest 40% were female. In table 4.8 are shown the results of the second evaluation process.

Table 4.8 The Weights of the Subdimensions of Integrated Index Version 2

The subdimensions of integrated index	The weights of each subdimension
V1 (Q5)	0.06
V2 (Q6)	0.87
V3 (Q7)	0.07
AE1 (Q8)	0.01
AE2 (Q9)	0.77
AE3 (Q10)	0.22
CM1 (Q11)	0.07
CM2 (Q12)	0.93
I (Q13)	1
CF1 (Q14)	0.26
CF2 (Q15)	0.18
CF3 (Q16)	0.06
CF4 (Q17)	0.50
FH1 (Q18)	0.38
FH2 (Q19)	0.13
FH3 (Q20)	0.48
M1 (Q21)	0.22
M2 (Q22)	0.10
M3 (Q23)	0.62
M4 (Q24)	0.06
COM1(Q25)	0.42
COM2 (Q26)	0.58
AC1 (Q27)	0.32
AC2 (Q28)	0.52
AC3 (Q29)	0.16
FL1 (Q30)	0.63
FL2 (Q31)	0.37
RE1 (Q32)	0.23
RE2 (Q33)	0.19
RE3 (Q34)	0.58
EP1 (Q35)	0.82
EP2 (Q36)	0.16
EP3 (Q37)	0.02
CS1 (Q38)	0.15
CS2 (Q39)	0.22
CS3 (Q40)	0.63
VLT1 (Q41)	0.72
VLT2 (Q42)	0.28
AS1 (Q43)	0.09
AS2 (Q44)	0.91
Integrated Index Score	0.87

Similarly, by runnig PCA in Matlab we found out the subdimension and dimension weights and the Usability Index Score as presented in table 4.9 and in table 4.10.

Table 4.9 The Weights of the Subdimensions of Usability Index Version 2

The subdimensions of Usability Index	The weights of the subdimensions
Efficiency1 (Q1)	0.53
Efficiency2 (Q2)	0.47
Effectiveness1 (Q3)	0.43
Effectiveness2 (Q4)	0.57
Satisfaction (Q45)	1

Table 4.10 The Weights of the Dimensions of Usability Index Version 2

The dimensions of Usability Index	The weights of the dimensions
Efficiency	0.24
Effectiveness	0.40
Satisfaction	0.36
Usability Index Score	0.62

Comparing with the first version's integrated and usability index scores, in the second version of cell biology course we met with higher indexes. Usability index $\text{score}_1=0.59<\text{Usability index score}_2=0.62$ and Integrated index $\text{score}_1=0.69<\text{Integrated index score}_2=0.87$.

In fact, our UseLearn model proposes to stop the algorithm if one of the index scores is found higher after the improvements, we again wanted to run the algorithm once more for the second dataset.

To find out the weights of the dimensions of integrated index, we applied NN-UseLearn model using *feed-forward backpropagation* as network type, *trainlm* for training function and as transfer function in layer 1 *tansig* and in layer 2 *purelin*.

Table 4.11 The Summary of NN Models in UseLearn Case Study Version 2

Number of Neurons	MSE of Training Data	MSE of Testing Data	Epoch
3	3.8373E–004	0.0018	86
4	3.8376E–004	0.0017	100
5	3.8373E–004	0.0017	96
6	3.8373E–004	0.0022	63
7	5.6546E–004	0.0016	100
8	3.8373E–004	0.0017	88
9	3.8373E–004	0.0019	37
10	3.8373E–004	0.0026	12
11	3.8373E–004	0.0016	100
12	3.8373E–004	0.0020	29
13	3.8374E–004	0.0025	100
14	3.8373E–004	0.0023	25

Depending on the training and testing data errors as shown in table 4.11, for the second dataset the best fitting NN-UseLearn model was with 11 neurons in the first layer. So we can summarize our NN-UseLearn model as shown in figure 4.22.

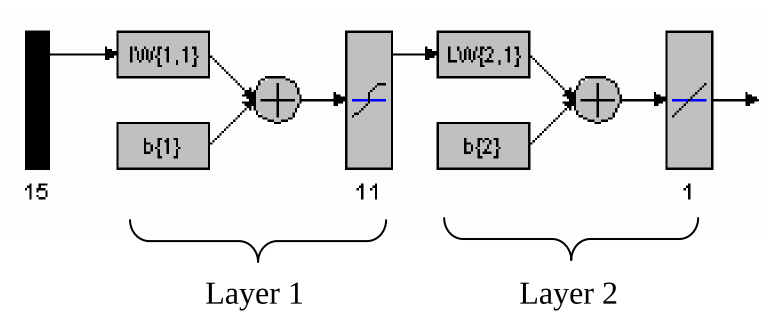

Figure 4.22 The View of the NN UseLearn Model for the Case Study Version 2

The model parameters of this NN-Uselearn model with 11 neurons are given in figure 4.22. As the figure 4.22 and figure 4.23 indicate, we randomly assigned one fourth of the data (18 observations) as testing data and three fourth of it (54 observations) as training data set.

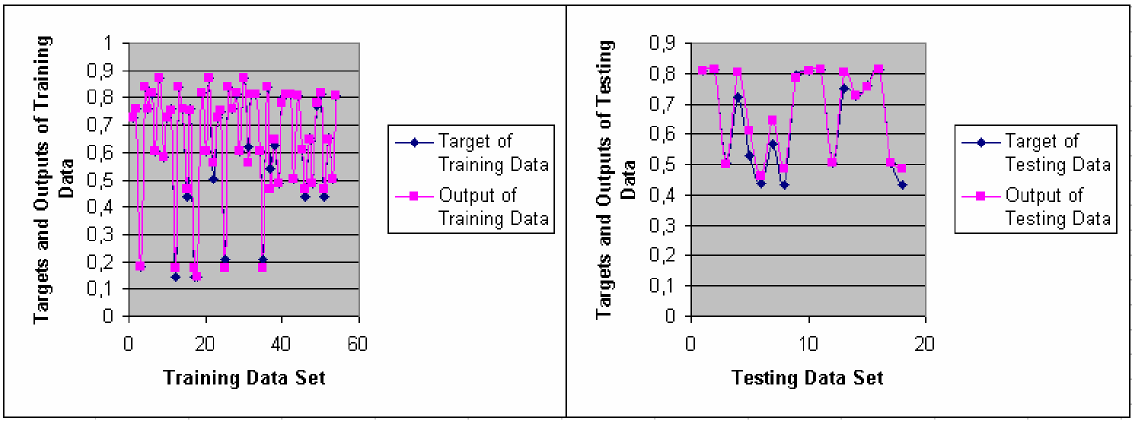

Figure 4.23 Targets and Outputs of Training and Testing Data Set Version 2

There are only e few outliers in the training data errors set and testing data set errors are fluctuating around zero, which means the model fits and is accepted to be satisfactory. These results are shown in table 4.24.

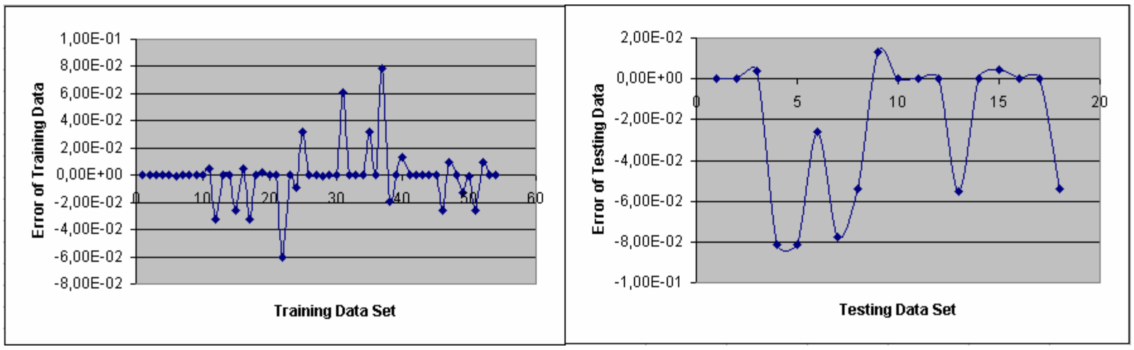

Figure 4.24 Errors of Training and Testing Data Set Version 2

A screenshot taken from Matlab shows the error depending on the number of neurons. As seen in figure 4.25, the lowest error is taken with 11 neurons.

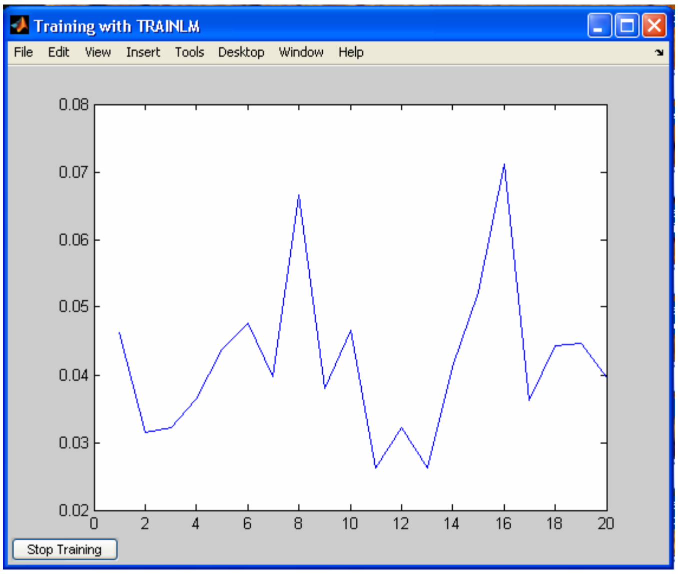

Figure 4.25 Matlab Screenshot of NN Model of Version 2

Another screenshot of Matlab represents the changing graph of *trainlm* training function as shown in figure 4.26.

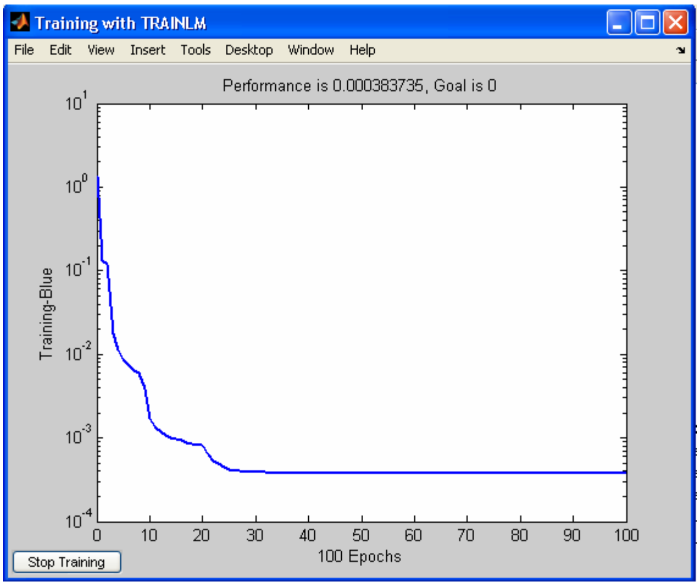

Figure 4.26 A View of Performance and Number of Epoch of trainlm Function Version 2

The weights of the dimensions found with this model are shown in table 4.12. We had to find out the most critical dimensions using these weights if we are eager to find out the usability problems, reprogram by improving these problems. As done in the first version here we again applied Pareto Analysis.

Table 4.12 The Weights of the Dimensions of Integrated Index Version 2

The abbreviation of the dimensions	The weights of the dimensions by NN
V	0.053
AE	0.084
CM	0.071
I	0.013
CF	0.034
FH	0.205
M	0.007
COM	0.034
AC	0.097
FL	0.007
RE	0.048
EP	0.060
CS	0.077
VLT	0.098
AS	0.112

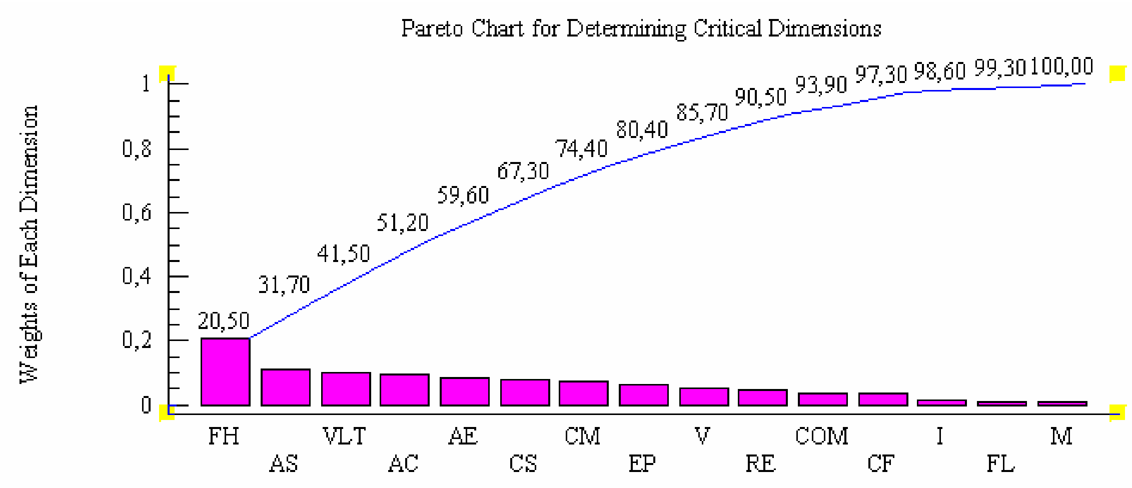

Figure 4.27 Pareto Analysis for Determining Critical Dimensions Version 2

If Pareto Analysis applied to these weights, the figure 4.27 is reached. If the usability evaluater expert is still determined to do improvements, he has a lot of tasks to do this time. This is because first time the integrated index score was calculated as 0.69 but in the second version the integrated index score was 0.87. So to get a higher integrated index score, it is needed to tackle with much more dimensions. In the first version it was enough to tackle with 33% of the dimensions to eliminate 78.30% of the problems. But this time depending on the

Pareto Analysis in figure 4.27, it is needed to focus on (8/15=0.53) 53% of the dimensions to eliminate 80.40% of the usability problems. These dimensions are *Feedback&Help, Assessment Strategy, Vareity of Learning Tasks, Accessibility, Aesthetics, Clear Syllabus, Course Management, and Error Prevention.*

4.3.2 Usability Problems and Design Recommendations

Next step is to find out to handle with which subdimensions. For each dimension, we can focus on the subdimensions whose sum of weights (or only weight) made up 20 % or more just to go on our hypothesis related to Pareto Analysis defining the threshold value. Take care of that the weights in parenthesis are 20 % or more:

FH3: Is the user provided with sufficient information to know where in the system he/she is?(0.48)

AS2: Are there multiple assessment strategies to measure content knowledge, skills, and performance standards? (0.91)

VLT1: Does the course provide variety of visual, textual, and/or auditory activities? (0.72)

AC2: Are alternative pathways to course content and activities available (0.52)

AE2: Is there sufficient use of white space in the course design? On each page? Across course pages? (0.77)

CS3: Does the course syllabus identify and clearly delineate the role that the online environment will play in the total course? (0.63)

CM2: Does the course provide specific resources to support online student learning? (0.93)

EP1: Can the user easily undo selections, actions, errors in arrangement or management of items? (0.82)

Note that *Course Management* and *Error Prevention* dimensions are the same usability problems that took place in the first version. But the subdimensions of them changed depending on the weights of PCA algorithm. In the first version, the subdimensions expected to be improved were EP3 and CM1. In this version EP1 and CM2 seem to be more critical because the previous ones were improved.

4.4 SUMMARY OF RESULTS OF CHAPTER 4

In chapter 4, *UseLearn* was implemented by the help of an online cell biology course using the eLearning system, *Moodle*. A case study for usability evaluation of eLearning system in cell biology by UseLearn at Fatih University and Fatih College is carried out. At the first stage, data was collected from 88 students. By UseLearn-based analysis usability problems were allocated and relevant design recommendations were defined. So the eLearning system was improved. At the second stage, data was collected from 72 students by using the improved eLearning system. The analysis of experimental results showed that UseLearn approach supports allocation of usability problems and defining relevant improvement measures. The main advantage of UseLearn is the adaptive selection of most significant checklist dimensions and thus reducing the time and effort for usability evaluation. So we can claim that *UseLearn* algoritm is the foremost quantitative method of usability testing methods because the results of the case study proved that this algorithm is succeeding to recover the usability problems and improving the usability satisfaction. This recovery is achieved by just focusing on some of the usability problems not all and the error rates of neural networks models were also satisfactory.

CHAPTER 5

CONCLUSIONS

In chapter 2, traditional learning and eLearning are compared and the benefits of eLearning are defined. Usability evaluation and quality evaluation approaches of eLearning systems are compared. There are many overlapping items in these approaches depending on the natural result of interaction between usability and quality. There is evident need of developing a widely accepted quantitative method applicable to eLearning systems measuring the usability of eLearning systems and determining the usability problems.

In chapter 3, a checklist including both quality and usability dimensions was developed. Based on the checklist, a common quantitative adaptive usability evaluation method *UseLearn* was proposed. It includes many different quantitative techniques such as principal components analysis, neural networks and pareto analysis.

In chapter 4, *UseLearn* was implemented by the help of an online cell biology course using the eLearning system, *Moodle*. A case study for usability evaluation of eLearning system in cell biology by UseLearn at Fatih University and Fatih College is carried out. At the first stage, data was collected from 88 students. By UseLearn-based analysis usability problems were allocated and relevant design recommendations were defined. So the eLearning system was improved. At the second stage, data was collected from 72 students by using the improved eLearning system. The analysis of experimental results showed that UseLearn approach supports allocation of usability problems and defining relevant improvement measures. The main advantage of UseLearn is the adaptive selection of most significant checklist dimensions and thus reducing the time and effort for usability evaluation. So we can claim that *UseLearn* algoritm is the foremost quantitative method of usability testing methods because the results of the case study proved that this algorithm is succeeding to recover the usability problems and improving the usability satisfaction. This recovery is achieved by just focusing on some of the usability problems not all and the error rates of neural networks models were also satisfactory.

A further study for improving this study can include:

1. applying neural networks in both steps of finding out the weights of dimensions and subdimensions (Remember that we applied principal components analysis for finding the weights of the subdimensions and neural networks for finding the weights of the dimensions).

2. applying structural equation modeling and another adaptive neuro-fuzzy model, ANFIS to all steps and comparing the error rates of all these models and selecting the best fitting one.

3. another application of threshold value different than Pareto analysis for determining which dimensions to focus on. (Remember that we used in UseLearn algorithm Pareto analysis for determining threshold value).

REFERENCES

Ardito, C., M.D., Marsico, R., Lanzilotti, S., Levialdi, T., Roselli, V., Rossano, M., Tersigni, "Usability of eLearning Tools", *Gallipoli (LE), Italy*, 2004.

Barker, K.C., "Quality Standards for Consumer Protection", *Learning Innovations Forum*, August 2004.

Bates, A. W., "Technology, Open learning and Distance Education", *Routledge*, London, 1995.

Bergman, E. and R. Haitani, "Designing the Palm Pilot in Bergman, E. Information Appliances and Beyond", *Morgan Kaufmann: San Francisco*, 2000.

Bevan, N., "Usability is Quality of Use", *Proceedings of the 6th International Conference on Human Computer Interaction, Yokohama,. Anzai & Ogawa (eds), Elsevier*, July 1995

Clarke, A., "E-learning Skill", *Palgrave Macmillan p. 1-19*, 2004.

Daniel, J. S., "Mega-Universities and Knowledge Media: Technology Strategies for Higher Education", *Kogan Page*, London, 1996.

Dawson, P. and G. Palmer, "Quality management", *Longman*, Australia, 1995.

Dede, C., "The Evolution of Distance Education: Emerging Technologies and Distributed Learning", *The American Journal of Distance Education*, V: 10 (2):p.4-36, 1996.

Deming, W.E., "The new economics", *(2nd ed.) USA: Massachusetts Institute of Technology Centre for Advanced Educational Services*, 1994.

Dringus, L., "An Iterative Usability Evaluation Procedure for Interactive Online Courses", *Journal of Interactive Instruction Development, Vol. 7, N. 4, pp. 10-14*, 1995.

Dringus, L. And M. S. Cohen, "An adaptable Heuristic Checklist for Online Courses", *35th ASEE/IEEE Frontiers in Education Conference T2H-6, Indianapolis*, October 2005.

Euler, D. and S. Seufert, "Workshop eLearning Quality", *Swiss Center for Innovations in Learning, University of St. Gallen*, 2005.

Feldstein, M., "Ignore Usability At Your Peril", *eLearn Magazine"*, *http://www.elearnmag.com/subpage/sub_page.cfm?article_pk=5142&page_number_nb*, 2002.

Frokjaer, E., M., Hertzum and K. Hornbaek, "Measuring Usability: Are Effectiveness, Efficiency, and Satisfaction Really Correlated?", *CHI Papers 1-6 April 2000 Volume 2, Issue1*, 2000.

Frontend, A., "Why People Can't Use eLearning", *Whitepaper*, *http://infocentre.frontend.com/downloads/Why_people_can't_use_eLearning*, 2001

FuturED Inc. and the Canadian Association for Community Education, "CanREGS: Canadian Recommended eLearning Guidelines:Quality Standards", 2002.

Garvin, D. ,"Managing quality*", New York: Macmillan*, 1988.

Gilbert, J., "A slice by slice guide to total quality management", *Great Britain: Tudor*, 1992.

Gilmour, P. and R. Hunt, "Total quality management", *Australia: Longman*, 1995.

Grahan, C., Cagiltay, J. Lim, B., Craner, J., & Duffy, T.M., "Seven Principles of Effective Teaching: A Practical Lens for Evaluating Online Courses", *Michigan Virtual University*, 2002.

Hagar, P., "Quality assurance in vocational education and training", in C. Robinson and P. Thompson (Eds.), *Readings in Australian vocational training and training research* (pp.340-362),Australian National Training Authority, Australia, 1998.

Hiltz, R. H. ,"The Virtual Classroom: Learning Without Limits via Computer Networks", "*Norwood, NJ: Ablex"*, 1995.

ISO, "Ergonomic Requirements for Office Work With Visual Display Terminals (VDTs)-Part 11: Guidance on Usability", *(ISO 9241-11:1998(E)). Geneva, Switzerland*, 1998.

ISO 13407, "Human-centred design processes for interactive systems", 1999.

Jolliffe, I.T., Principal Component Analysis, New York: Springer-Verlag, 1986.

Krug, S., "Don't Make Me Think*", circle.com library, p. 5 Whitney Quesenbery, Cognetics Corporation*, 2000.

Kruse K., "eLearning and the Neglect of User Interface Design", *eLearningGuru.com* , *http://www.e-learningguru.com/articles/art4_1.htm*,1999.

López, A.J., Pérez, R., Fernández, M.M., Vicente, M.R., "Approaching The Quality Of The Spanish Universities Through ICT Indicators", *Paper to the European Society for Organizational Excellence (ESOE) conference on TQM for Higher Education Institutions*, 2003.

McIsaac, M. S. and C. N. Gunawardena, "Handbook of Research for Educational Communications and Technology", *ed. D. Jonassen, p. 403-437. New York: Simon and Schuster Macmillan*, 1996.

McNamara, P., "Usability and eLearning", *Certificate in Online Education and Training-Institute of Education University of London*, May 2003.

Miller, M.J., "Usability in eLearning", January 2005.

Moodle, "Open Sourced Course Management System for Online Learning", *www.moodle.org* and *http://moodle.fatih.edu.tr* .

Nielsen, J., "Usability Inspection Methods", *Wiley, New York*, 1994.

Nielsen,J.,"eLearningpost on eLearning", http://www.elearningpost.com/features/archives/001015.asp, 2001.

Moore, J. C., "The Sloan Consortium Quality Framework and The Five Pillars", 2004.

Moore, M. G. and G. Kearsley, "Distance Education: A System's View", *Wadsworth Publishing Co.*, 1996.

Nichols, M., "Development of A Quality Assurance System for eLearning Projects", *Ascilite Conference Proceedings*, 2002.

Notess, M., "Usability, User Experience, and Learner Experience", *http://www.elearnmag.org* , 2001

Pawlowski, J.M. and U. Ehlers, "CEN/ISSS Workshop Learning Technologies: Quality Standards for E-Learning", *Information Systems for Production and Operations Management,Helsinki*, June 2005.

Piguet, A. and D. Peraya, "Creating web-integrated learning environments: An analysis of WebCT authoring tools in respect to usability", *Australian Journal of Educational Technology, 16(3), 302-314, http://www.ascilite.org.au/ajet/ajet16/piguet.html* , 2000

Quality Assessment Rubric for eLearning Design, NetTel@Africa Course, Panel of eLearning Experts.

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

Reeves, T.C., "Evaluating What Really Matters in Computer-Based Education", *University of Georgia*, 1997.

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

Schwier, R.A. and E.R. Misanchunk, "Interactive Multimedia Instruction", *Englewood Cliffs, NJ: Educational Technology Publications*, 1993.

Smith, P. L. and C. L. Dillon, “Comparing Distance Learning and Classroom Learning: Conceptual Considerations”, *The American Journal of Distance Education,* No: 13: p. 6-24, 1999.

Smulders, D., “Designing for Learners, Designing for Users, eLearn Magazine”, *http://www.elearnmag.org/subpage/sub_page.cfm?section=3&list_item=11&page=1.*

Squires, D. and J. Preece, “Predicting quality in educational software: Evaluating for learning, usability and the synergy between them”, *Interacting with Computers 11(5):* 467–483, 1999.

STC Conference Proceedings, “What Does Usability Mean: Looking Beyond ‘Ease of Use’ ”, *by Quesenbery, Whitney,* 2001.

Stein, G., “Putting the e into Learning & Teaching: Quality Assurance Guidelines”, *Client-Communications & Information Technology, Canterbury Christ Church University, UK.*

Stephenson, J., “Definitions of Indicators of Quality on the Application of ICT to University Teaching”, *Paper Workshop at Tarragona*, Spain, September 2005.

Villiers, R.D., “Usability Evaluation Of an E-Learning Tutorial: Criteria, Questions and Case Study”, *Proceedings of SAICSIT, Pages 284 – 291,* 2004.

Vogl, T. P., J.K. Mangis, A.K. Rigler, W.T. Zink, and D.L. Alkon, "Accelerating the convergence of the backpropagation method," *Biological Cybernetics, vol. 59, pp. 257-263*, 1988.

Walklin, L., “Putting Quality into Practice”, *England: Stanley Thornes Ltd.*, 1992.

Whitworth, J. M., “Looking at Distance Learning through Both Ends of the Camera”, *The American Journal of Distance Education,* No: 13(2): p. 64-73, 1999.

Wilkinson, T., “Summer Enrollment in Distance and Distributed Learning Courses”, *The Institute for Distance and Distributed Learning Virginia Tech.*, September 2005.

APPENDIX A

PRINCIPAL COMPONENTS ANALYSIS FOR VERSION 1

PCA Scores for *Visibility*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,9257	64,190	64,190
2	0,703102	23,437	87,627
3	0,371199	12,373	100,000

PCA Scores for *Aesthetics*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,64731	54,910	54,910
2	0,709328	23,644	78,555
3	0,643359	21,445	100,000

PCA Scores for *Course Management*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,76441	88,220	88,220
2	0,235594	11,780	100,000

PCA Scores for *Consistency and Functionality*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	2,01363	50,341	50,341
2	0,907298	22,682	73,023
3	0,587458	14,686	87,710
4	0,491616	12,290	100,000

PCA Scores for *Feedback and Help*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,6584	55,280	55,280
2	0,786266	26,209	81,489
3	0,555332	18,511	100,000

PCA Scores for *Memorability*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	2,34382	58,595	58,595
2	0,7389	18,472	77,068
3	0,571583	14,290	91,358
4	0,3457	8,642	100,000

PCA Scores for *Completeness*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,5472	77,360	77,360
2	0,452796	22,640	100,000

PCA Scores for *Accessibility*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,53161	51,054	51,054
2	0,918363	30,612	81,666
3	0,550031	18,334	100,000

PCA Scores for *Flexibility*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,00978	50,489	50,489
2	0,990221	49,511	100,000

PCA Scores for *Reducing Redundancy*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,69285	56,428	56,428
2	0,803721	26,791	83,219
3	0,503429	16,781	100,000

PCA Scores for *Error Prevention*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,76597	58,866	58,866
2	0,750532	25,018	83,884
3	0,483493	16,116	100,000

PCA Scores for *Clear Syllabus*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	2,24679	74,893	74,893
2	0,452619	15,087	89,980
3	0,300594	10,020	100,000

PCA Scores for *Variety of Learning Tasks*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,55492	77,746	77,746
2	0,445082	22,254	100,000

PCA Scores for *Assessment Strategy*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,30236	65,118	65,118
2	0,697641	34,882	100,000

APPENDIX B

PRINCIPAL COMPONENTS ANALYSIS FOR VERSION 2

PCA Scores for *Visibility*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,50047	50,016	50,016
2	0,963683	32,123	82,139
3	0,535844	17,861	100,000

PCA Scores for *Aesthetics*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,76597	58,866	58,866
2	0,750532	25,018	83,884
3	0,483493	16,116	100,000

PCA Scores for *Course Management*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,2546	41,820	41,820
2	0,881288	29,376	71,196
3	0,864112	28,804	100,000

PCA Scores for *Consistency and Functionality*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	2,34382	58,595	58,595
2	0,7389	18,472	77,068
3	0,571583	14,290	91,358
4	0,3457	8,642	100,000

PCA Scores for *Feedback and Help*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,25605	41,868	41,868
2	0,929459	30,982	72,850
3	0,814491	27,150	100,000

PCA Scores for *Memorability*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,58903	39,726	39,726
2	0,969307	24,233	63,958
3	0,860359	21,509	85,467
4	0,581308	14,533	100,000

PCA Scores for *Completeness*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,1353	56,765	56,765
2	0,864699	43,235	100,000

PCA Scores for *Accessibility*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,63734	54,578	54,578
2	0,7694	25,647	80,225
3	0,593262	19,775	100,000

PCA Scores for *Flexibility*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,41792	70,896	70,896
2	0,582078	29,104	100,000

PCA Scores for *Reducing Redundancy*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,59378	53,126	53,126
2	0,914378	30,479	83,605
3	0,491847	16,395	100,000

PCA Scores for *Error Prevention*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,13577	37,859	37,859
2	0,98603	32,868	70,727
3	0,878204	29,273	100,000

PCA Scores for *Clear Syllabus*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,1688	58,440	58,440
2	0,831196	41,560	100,000

PCA Scores for *Variety of Learning Tasks*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,34191	67,095	67,095
2	0,658093	32,905	100,000

PCA Scores for *Assessment Strategy*

Principal Components Analysis

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1,06489	53,245	53,245
2	0,93511	46,755	100,000