A Measurement Model Based on Usability Metrics for Electronic Assessments (Case- Study, Faculty of Information Technology)

By

Eman Issa Saleh Belgasem

Supervisor

Dr. Kenz A. Bozed

Dissertation submitted to the Faculty of Information Technology in partial fulfillment of the requirement of Master‘s degree in Software Engineering

University of Benghazi

Faculty of Information Technology

2018 July

University of Benghazi **Faculty of Information Technology**

Department of Software engineering

A Measurement Model Based on Usability Metrics for Electronic Assessments (Case- Study, Faculty of Information Technology)

By
Eman Issa Saleh Belgasem

This Thesis was Successfully Defended and Approved on **23.7.2018**

Supervisor
Dr. Kenz Amhmed Bozed

Signature: **……..**

Dr. Omar Mustafa Elsalabi (Internal examiner)

Signature: ……..

Dr. Abdelhamead M. Abdelkafy (External examiner)

Signature: ……**.....**

(Dean of Faculty) **(Director of Graduate studies and training)**

In the name of Allah clement, the philanthropist, owner of the King.

Dedications

For: my parents, my dear brother Saad, and my sisters Aisha, Asma, Masouda, Nowara, our grandchildren: Jana, Moaad, and the twin of my soul Walid for their deep love, encouragement, and patience.

Eman

Acknowledgment

"The smell of perfume will still, at the hand that gives you the roses"

Praise and thank to Allah who gave me the determination to complete this dissertation, then many thanks go to my supervisor Dr. Kenz Bouzid Hassouni for her support, encouragement, and guidance to complete my dissertation .

I extend my thanks and deep appreciation to Dr. Ashraf Huwaidi, Dr. Saleh AL Shalmane and Dr. Mohamed Eklaif for supporting me during my scientific studying; also, I do not forget my tutors and friends who accompanied me through my scientific studying.

All thanks, because thanks will make deep respect.

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

LIST OF ABBREVIATIONS

CBAs	Computer Based Assessments
E-assessments	Electronic assessments
HCI	Human Computer Interaction
ICT	Information and Communication Technology
ISO	International Organization of Standardization
IT	Information Technology Faculty
JISC	Joint Information System Committee
MMB-UM	A Measurement Model Based on Usability Metrics
PBAs	Paper Based Assessments
UX	User Experience

A Measurement Model Based on Usability Metrics for Electronic Assessments (Case- Study, Faculty of Information Technology)

By

Eman Issa Saleh Bilqasim

Supervisor
Dr. Kenz Amhmed Bozed

Abstract

The modern scientific revolution influences at the educational system, especially at critical part of education process which is(the assessment), the traditional assessment is consists some complicated issues relate to the time, the effort, in addition, to the consumption cost of the papers. The Electronic Assessments(E-Assessments) is considered one application of Human-Computer Interaction (HCI). Therefore, can be measured Usability and User Experience(UX) for E-assessments, in this dissertation, applied the proposed prototype of E-assessments for the first time in the Faculty of Information Technology at Benghazi University.Conducted an E-assessment experiment on a group of the students and some tutors,Online assessment and Computer Based Assessments (CBAs) offline assessment, then compared the results between them and measured the most important factors usability and UX by the Measurement Model Based Usability Metrics(MMB-UM) for E-assessment.

Chapter1

Introduction

1.0 Overview

This research is an empirical investigation into Usability and User Experience (UX), which are considered Human-Computer Interaction (HCI) perspectives in Electronic Assessments (E-Assessments) by using A Measurement Model Based on Usability Metrics (MMB-UM). This chapter will provide a brief background to the topic and discuss problem statements, motivation, aims, and objectives of the research.

1.1 Background

In the midst of acceleration of the life and the changes, which brought by the information technology. The people look for something to make their lives easy and enjoyable, that led to the question, which asked by Donald (1998, p. 28-30) when he describes the contradiction of technology *"But what good is technology if it is very complicated to use."* This question made a watershed in the field of Human-Computer Interaction (HCI) that plays an indispensable role in the development of the computer applications. Carroll (2002) defined HCI as a: " very important concept in the system development process as it is about understanding and creating software and other technology that people will want to use will be able to use, and will find effective when used. Also, the Usability concept and the methods and tools to encourage it, achieve it, and measure it are now touchstones in the culture of computing". In addition, Plaomba and Banta (1999) cited Usability as a concept appeared as a result of HCI and it is considered one of the most important quality factors for the web application. Also, (ISO9241-11, 1998) defines Usability as:*" The extent to which a product can be used by specified goals with effectiveness, efficiency, and satisfaction in the specified context of use."* Recently, Rajanen and Clemmesen, etc. (2017) contributed and indicated there is no a circumscribed definition for Usability and UX. Therefore, these concepts have influence in the field of education, then through the emergence of software applications, and sites, which supporting tutors and students to reach the desired precision; also, enhance of the education process level. Educational

evaluation of students considered an essential element in determining the level of the student and the success of the education process. Electronic-Assessments (E-Assessments) are offering a new idea to evaluate students by using concepts and tools, to help the tutors in the educational process. For an instant, Computer -Based Assessment (CBAs) and Online tests have become a widespread and growing in use, which are require professionalism to test all types of the students, especially those with disabilities (e.g. Thelwall 2000, Matraf and Hussain 2017). Similar to the Universities in developed countries, which have begun using CBAs in assessing students. The goal of this research is to introduce A Measurement Model to evaluate Usability and UX (User Satisfaction) for the proposed prototype of the CBAs at Benghazi University.

1.2 The Problem Statement

The reason for this dissertation is the challenges and problems which facing the tutors in assessing, students from all aspects are: the time, the effort, and cost; in addition, to the phenomenon of the cheating, and the traditional methods in student's assessment at Benghazi University. Indeed, the Information Technology Faculty (IT) should be studying the new and trendy technology and rivalry in the current development in the world. This development consists of the student's assessment process as well. Where the E-Assessments had been used as an alternative to the Paper Based Assessments (PBAs). Therefore, Usability and UX are important factors to measure the quality of E-Assessments; there is overlap between Usability, and UX need to be clarified.

This has led to the question of the dissertation is: *How is designing E-Assessments, efficiency and efficacy and safety evaluation, and obtain satisfaction tutors and students?*.

1.3 Motivation

Most universities in developed countries tended towards the use of E-Assessments, and in the Middle East some Arab universities began to apply this CBAs, for example : Yarmouk University in Jordon, Abdul Aziz University, which is considered a leader in the application of electronic tests in Saudi Arabia Universities, which began to implement the tests electronic starting from the academic year (2009/2010), and since then, the university continuous the process in the expansion and development , where

it was one of the most important stages for the purpose of development to upgrade to the E-assessments (Question mark Universal), which will be launched during the school year (2014/2015). (www.King Abdul Aziz University).This dissertation is to contact with the reality of the assessment for students at the Information Technology Faculty from the standpoint of tutors and students. Therefore, E-Assessments will be introduced and applied for the first time in the Faculty of Information Technology (IT) - University of Benghazi. Also, studying the feedback from the users who will run the experiments and find out their opinions, their satisfaction with the proposed prototype, and measure HCI perspectives (Usability, user experience).

1.4 Research aims and objectives

The aims are:

1. Design a proposed prototype CBAs to assessment the students of IT Faculty University of Benghazi.
2. Evaluate the proposed prototype CBAs by measuring factors in the HCI perspectives by introducing A Measurement Model Based on Usability Metrics (MMB-UM), in order to evaluate the implemented proposed prototype CBAs at IT Faculty.

The objectives are:

1. Propose a prototype CBAs of evaluation for the students in IT.
2. Focus on the requirements of the tutors and the students' aspirations, whereas the tutor's requirement developing the tools for saving the time, the effort and the cost when they are evaluating students, and the students desire to do their exams in a scientific environment comfortable.
3. Illustrate the perspectives of HCI in the concept of Usability and User Experience (UX) and the differences between them.
4. Reduce the phenomenon of cheating by making different versions of the tests, where the computer will select at random a copy exam for each student.

1.5 Research methodology

The research methodology central to this work is as follows:

1. Literature Review: Investigate and summarize previous researchers' work in relation to the CBAs and perspectives of HCI.
2. Problem Definition and Scope: Study and find out the relevance techniques of existing work in order to define the problem of the research aims and objectives.
3. Proposed Prototype Design and Implementation:
    - Design the proposed prototype of CBAs in formulating standards, which able to evaluate students in successful.
    - Implement the proposed prototype by the suitable tool.
    - Obtain the feedback from the students and the tutors.
4. Experimental Results: Experiments are performed on the proposed prototype of CBAs assessments and online assessment for the students using A Measurement Model Based on Usability Metrics (MMB-UM).
5. Proposed Prototype of CBAs Testing / Evaluating: Test the proposed prototype and evaluate the functionality (Usability and UX) compared the proposed prototype of CBAs with online assessment results.

1.6 Scope and limitations of the research

1. The research confined to students of Information Technology Faculty, at Benghazi University.
2. Concentration on the proposed prototype E-Assessments for undergraduate students.
3. The E-Assessments questions were used multiple choices.
4. The experiments were conducted in the confine lab.

1.7 Dissertation Organization

The dissertation was organized as the following:

Chapter 2: Review background material that includes knowledge on Human-Computer Interaction (HCI), Computer- Based Assessments (CBAs) and the relevance of existing works of this research.

Chapter 3: The methodology used for this research, in terms of identifying participants and how to collect data, selecting the appropriate tools for the design of the proposed CBAs model, the appropriate site for the establishment of online Assessments, and the overall implementation of the methodology

Chapter 4: Implement the proposed prototype of CBAs and perform the experiments to compare between CBAs and online assessment. After that, measure Usability and UX in order validate Measurement Model Based on Usability Metrics (MMB-UM), and discuss experimental results in details.

Chapter 5: Introduce the conclusion, contributions, and the recommendations for future work.

Chapter2

Literature Review

2.0 Introduction

This chapter will give an overview of the literature review that focuses on Electronic Assessments (E-Assessments) and Human-Computer Interaction (HCI) it is organized as follows:-it provides a definition of E-Assessments explore its applications and tools; also, advantages and disadvantages E-Assessments, then the definition of the HCI perspectives, and relevance between the previous research and this research.

2.1 Human - Computer Interaction (HCI)

The interface is methodical that the human begins to interact with computers, such as what is displayed on the computer screen, keyboard, and other equipment. HCI study emerged as a result of that, it is focusing on the satisfaction of users; because bad interface leads to the failure of the best product. HCI is *"This term acknowledged that the focus of interest was bolder than just the design of the interface and was concerned with all those aspects that relate to the interaction between users and computers"*Preece et al (1994, p. 7).

HCI emerged in the early 1980s in the last century; it was considered a broad area of research and practice. It started as a specialized area within the computer science; nestled cognitive science and engineering human factors. HCI has expanded very rapidly and incrementally over three decades; which attracted the experts from multiple disciplines and in overlap in many concepts.

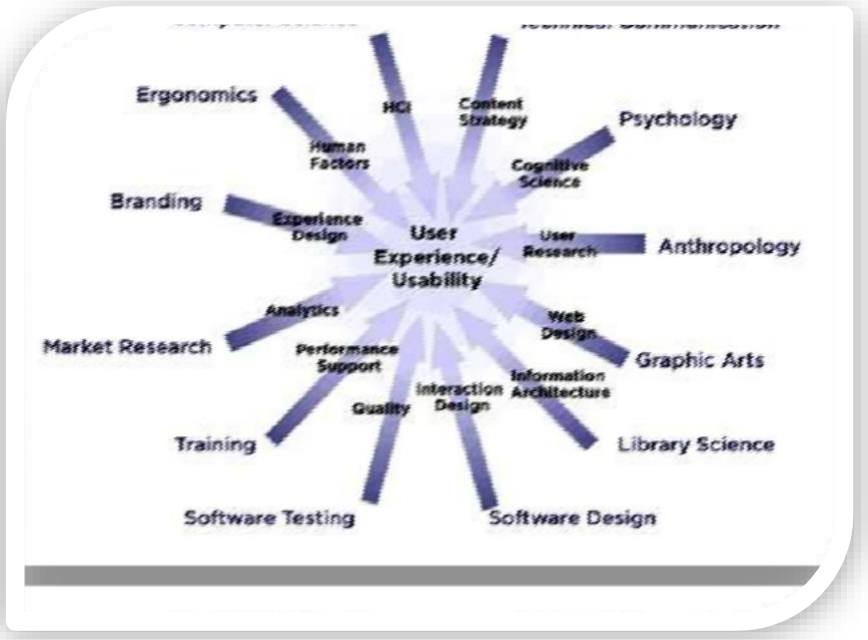


Figure 2-1: User experience professionals come from many backgrounds.

Source: (Quesen berry, 2011) © UPA, found on the UPA website.

The main objective of the HCI researches is designing a computer proposed prototype; to help users to end the necessary functions effectively. HCI plays an important role in the development of computer proposed prototypes as it helps to develop *" instructional techniques and to suggest where and in what situations ,these technologies and techniques might be put to best use."* Booth (1989, p. 6). The computer in the beginning of his appearance was a machine trying to simulate the human beings, but with the advances in technology it need human to communicate with a computer, this contact called HCI includes human interaction with the computer in the presence of hardware, software, this is the so-called interactive system Sim and Gavin (2009).For example Plantak Vakovac ,et al (2010) cited that: the field of human-computer interaction (HCI) interest in Usability and UX issues; whereby HCI is “how to work with and improve the Usability of interactive System" Hornbæk (2006) .

2.1.1 The development of HCI

Scientists through their researches classified HCI development, According to Yang and Chen (2009) divided into three stages as follows:

- Founding period (1970-1973): The concept was created.
- Developing period (the 1980s): HCI developed gradually.
- Improving period (the 1990s): Intelligent interactive; "users are part of it ."

Therefore, by based on all previous concepts and stages of HCI and according to Rajanen and Clemmensen (2017) the Usability concept appeared as a consequence of the HCI. The concept of Usability appeared as consequence within the field of human-computer interaction; because the users need more than just an interactive proposed prototype, also they need (fun, aesthetic, etc.), all of that lead to call the Usability: is the quality of the full understanding of the principle of HCI. To all HCI design, Usability and UX goal awareness is essentially as follows:

Usability: *"Through specific computer system criteria, include: efficiency, safety, utility, and learning/retention is "a central to interaction design and operations."*

UX: *Is "satisfying, enjoyable, entertaining, helpful, motivating aesthetically pleasing, creativity supportive, rewarding, fun and emotionally fulfilling "*concentrating on creating enjoyable and integrated interactive system Carroll (2002).

2.2 Computer - Based Assessment (CBAs)

The developments in computer technology and informatics, insert many new methods for education, to change is the life of students and the tutors. One of these methods to use computers in teaching and testing students since of the 80s based on according to Brown (1997). The computer-based assessment can contain elements of the multi-media such as video, images, and sounds, different from the limited evaluation over PBAs Morgan and Spector (2004) lists five methods of collecting students answer in CBAs:

1. Choose the answer.
2. Write short answer questions.
3. The arrangement and choice.
4. The connection between the answer and questions symmetry.
5. Locate the image and write the correct answer on the image.

(JISC, 2007:6) as cited in Nottingham Trent University(2016) : the CBAs distinguishes from Computer Assisted Assessment (CAA) in doing all assessment process parts are delivered and corrected by computers; whilst CAA is only used as part of the assessment process. Features of the CBAs that can be used for diagnostic, formative, or summative assessment, also can be supervised or not _supervised.

Many researchers and international miens try to define" assessment" word; definitions of assessment in appendices B, According to Lynch (2002) *the purpose of assessment is making decisions or judgments about the students, by the system of collecting the information.* Form all above inference that: Assessment is the process, which tries to evaluate the knowledge, understand and skills, leaner owns it.

The performance of the educational system is affecting the global economy, which posed the technological innovation, attempted to adapt the changing in requirements of new technology. In the traditional learning environments; digital tools and technology are becoming standard, in many different environments, CBAs is already used widely. In essence, CBAs are a practice of giving quizzes and tests on the computer instead of using pencil and paper www.ProProfs Training Maker (2015).

2.2.1 Overview Computer-Based Assessment (CBAs) applications/tools

Based on the definition of (JISC, 2007:6) CBAs can be used for all assessment types summative**,** formative, or diagnostic**.** Typically, CBAs are based on multiple choice questions (MCQs) or other "objective" question types, although non-objective questions (e.g. Essay, short answer) can also be included.

Applications of CBAs: **Diagnostic**, **Formative** or **Summative** can be used by CBAs.

1. Diagnostic assessment

CBAs can be used with a diagnostic focus at the start of a course or module, or at other key points.This allows a tutor to evaluate students' baseline knowledge/ understanding and, if appropriate, provide feedback. It can also be used by staff to help identify gaps in student understanding at a cohort level.

2. Formative assessment

CBAs can be used to provide formative assessment, the i.e. assessment which is developmental for students' learning, but does not contribute to credit points. This can be done in-class or as a directed study. CBAs allows for instant feedback; this can be written to address student misconceptions directly, or to point students to relevant resources to support their own learning.

3. Summative assessment

CBAS can be used to provide summative assessment, i.e. assessment is that which results in a final grade (and feedback) which reflects the standard of achievement of the student work against intended learning outcomes. Summative CBAs may be especially useful for large cohorts, or in distance learning, where "economies of scale" are most applicable. A broader range of scenarios can be presented that would be possible in a paper-based exam, for example by including videos and simulations. Nature of traditional assessments is taking time at implementation; whence (preparation, monitoring, correcting, etc.), the exam summative assessment need for high effort exists invigilation and administration. Important questions that should be asked at design assessments are:

How will you verify students identify and prevent cheating? Are spare PCs available? Are contingency plans in place should there be a proposed prototype failure? Have you considered special access arrangements (e.g. For disabled students, or those with dyslexia or visual impairments?).

In recent years, increased the interest in developing of CBAs. Many commercial tools, freeware, and shareware tools; appeared as the result of studies and research in this field, that, and public institutions. Valenti , et al(2002) they are important to note that a typical CBAs system is composed by:

• A Test Management System (TMS) - i.e. a tool providing the instructor with an easy to use interface, the ability to create questions and to assemble them into tests, the possibility of grading the tests and making some statistical evaluations of the results.

• A Test Delivery System (TDS) - i.e. a tool for the delivery of tests to the students. The tool may be used to deliver tests using paper and pencil, a stand-alone computer, on a LAN, or over the web. The TDS may be augmented with a web-enabled used to deliver the tests over the Web. In many cases, producers distribute two different versions of the same TDS, one to deliver tests either on single computers or on a LAN, and the other to deliver tests over the web.

2.2.2 Advantages and disadvantages of CBAs

Overall, CBAs have main advantages discovered by the tutors as well as a few disadvantages, of using this technology.

One of the reasons that make the CBAs Best of PBAs is no need for the existence of the supervisors of the students during the tests, at the time of the exam in "some cases"; this gives a feeling of satisfaction for students. Formative or Summative assessments could be performed inside or outside the University; taking into account the instructions and restrictions of time, place and type of exam and the use of internet and intranet.

Despite this, there are some advantages and disadvantages of the use of online assessments and CBAs. By based on all the papers that have read; the sites which have visited; by the reference to the Simin and Heidari(2013); also the experiments that had been executing it on students through the research, I concluded the following:

Advantages of Computer Based Assessments (CBAs)

For the Students:

1. Depending on the interface that displays the questions will be the degree of satisfaction of the students.

2. The existence of forms of multi-media and interactive interfaces that make it easier for the students.
3. According to the regime, if provide the feedback to the students, to estimate the performance them in the exam.

The features of the tutors:

1. There is no need for the presence of observers on the students during the exams.
2. Reducing or eliminating the phenomenon of cheating, according to the time sand spatial constraints.
3. Get statistics and evaluations of the student without the need for the consumption of the timed effort. Low-cost (number of exam papers equal zero).
4. Time keeping the tutors can be done weekly examinations online.

Disadvantages of Computer Based Assessments (CBAs)

Disadvantages for students:

1. If there a specific time to answer the question; especially if it was the first time to interact student with the system.

Disadvantages for tutors:

1. The possibility of spoofing, especially online.
2. The exams only MSQ examinations which are not sufficient to test students scientifically.
3. The difficulty of CBAs implementation requires the cost may be higher in some cases.
4. May need to staff programming and staff to insert exams questions.
5. With reference to the Simth and Broom (2003), the students and the tutors still lack the skills to deal with the CBAs.

There are several disadvantages and advantages of CBAs. More details are enclosed in Appendices B.

2.3 Human - Computer Interaction (HCI) perspectives

HCI is performing the fundamental role in improving applications and websites, it will assist designers, analysts, and others identify the system needs, while Usability provides proposed prototype satisfied for the users. It is considered one of the most important factors to measure application quality. Alshamari and Mayhew (2009) believe that: *"Usability is one of the most important success factors in system quality"*. (ISO 9241-11, 1998) defines Usability as**:**" *The extent to which a product can be used by specified goals with effectiveness, efficiency, and satisfaction in the specified context of use"*. As discussed previously, the growth of the HCI has meant that the concept of Usability has become increasingly important; a user will view how usable a system is by whether they can achieve their goal Spool(2007).

Usability is relevant when discussing user experience: *"Usability, when interpreted from the perspective of the users' personal goals can include the kind of perceptual and emotional aspects typically associated with user experience*" ISO (2010), In an article by Spool (2007) discusses the difference between user experience and Usability, He believes that the term Usability asks the question, *"Can the user accomplish their goal?"*, while The question asked by the term user experience is *"Did the user have as delightful experience as possible?"* . Nilson(2003) presented to us the reasons why people will leave the application: *"(a) it is difficult to use; (b) if the users get lost on a website; (c) the information is hard to read; (d) it does not answer users' key quest*.

2.3.1 Usability

In the development process, Usability is considered a very important aspect, that it can mean the difference between performing and completing a task in a successful way without any frustration. Rajanen and Clemmensen (2017) they indicate that definition states Usability 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". Usability refers to the "*quality of the interaction in terms of*

parameters such as time taken to perform tasks, a number of errors made, and the time to become a competent user" Benyon et al(2005, p. 52). According to Nielsen (2003): *"Usability is a quality attribute that assesses how easy user interfaces are to use. The word "Usability" also refers to methods for improving ease-of-use during the design process."* There are varying sets of definitions specific attributes of Usability (facets, aspects, factors) defined by standards and authors. In Table 2-1 summarized some of these definitional sets. Concerning attributes of Usability apparent in the table row, for example, all sources in Table (2-1) describe "efficiency" as Usability attribute, although not all sources use this particular term (e.g. Nilson, 1993, Schneiderman, 1992).

Table 2-1: Sets of definitions specific attributes of Usability

Shackle (1991)	**Schneider man (1992)**	**Nielsen (1993)**	**Preece et al. (1994)**	**ISO 9241_11 (1998)**	**Constantine & Lockwood (1999)**	**Constantinos & Dan (2007)**
Effectiveness Speed	Speed of performance	Efficiency of use	Throughput	Efficiency	Efficiency in use	Efficiency
Learn ability (Ease of learning)	Time to learn	Learn ability (Ease of learning)	Learn ability (Ease of learning)	----------	Learn ability	----------
Learn ability (Retention)	Retention over time	Memorability	----------	----------	Remember ability	----------
Effectiveness (Errors)	Rate of errors by users	Errors/Safety	----------	----------	Reliability	Effectiveness
Attitude	Subjective Satisfaction	Satisfaction	Attitude	Satisfaction acceptability of use)	User satisfaction	Satisfaction

2.3.1.1 Overview of Usability Method

There are many techniques and metrics to measure the Usability in traditional applications, in which are some important epithets such as: learn_ ability, efficiency, memory -ability, error rate, and satisfaction Shawgi and Nureldine (2015). Over the last few years; Usability evaluation of software applications investigated widely over; Kabir and Han (2016) . According to Harrington(2017) that: *"Measurement is the first step that leads to control and eventually to improvement. If you cannot measure something, you cannot understand it. If you cannot understand it, you cannot control it. If you can't control it, you can't improve it"*. Generally, Usability can be evaluated by Inspection; Testing; Inquiry. Each part consists of parts in details:

1. **Usability Inspection Approach** may be; Usability specialists, software developers, users and other professionals examine Usability, related aspects of a user interface. Inspection Methods divide:
    - Heuristic Evaluation: Most common informal methods involve Usability specialist, they judge by basing on Usability principles.
    - Cognitive walk-through: involves one or a group of evaluators inspecting a user interface by going through a set of tasks and evaluate its understandability and ease of learning.
    - Action Analysis: Focus is on what the practitioner's do than on what they say they do focus is on.
2. **Usability Testing Approach** representative users work on typical tasks using the system and the evaluators use the results to see how the user interface supports the users to do their tasks, so testing methods include:
    - Coaching Method participants are allowed to ask any proposed prototype-related questions of an expert coach who will answer to the best of his or her ability.
    - Co-discovery Learning two test users attempt to perform tasks together while being observed. Whereas testing approach is:
    - Performance measurement this gives quantitative data, this test must be performed in Usability Laboratory because a small distraction can lead to false results.

- Question asking protocol: Users are asked direct questions about the product.
- Retrospective Testing: Videotape is made of the Usability test session. Tester reviews the videotape together with participants and asking questions about their behaviour.
- Teaching Method: Let the test user interaction with the system first. Let them be little expertise:
    1. Assign naïve user to each test user.
    2. Naive users are instructed not be given active participation in problem- solving.
    3. Test users are asked to verbalize their thoughts, feelings, and opinions while interacting with the system.
- Thinking Aloud Protocol: Test users are asked to verbalize their thoughts, feelings, and opinions while interacting with the system.

3. Usability Inquiry Approach

Field Observation; Focus Groups; Interviews; Questionnaire; Logging Actual Users; Proactive Field Study. More details enclosed in Appendices A.

2.3.1.2 Overview Usability Models

Usability concept has been defined in multiple ways, also evolved over a period. Holbrook et al (2006) coincidence that there are three major ISO standards which include: ISO 9241 -11 (1998) it identifies efficiency, effectiveness, and satisfaction as key attributes. ISO/IEC 9126-1 (2001) define Usability as software quality attribute decomposed into five factors understandability, learn-ability, operability, attractiveness, and Usability compliance. There are many Usability models, but no one of these models covers all aspects of Usability. Such as Usability, models include Nielsen (1993), Shneiderman (1992), Preece et al. (1994), Shackel (1991) and Constantine & Lockwood (1999).

Based on the above and the features of CBAS proposed prototype, proposed an improved Usability evaluation model with five Usability factors for CBAs. The details are listed in Table (2-2) and each factor corresponds to a description.

Table 2-2: Proposed Usability Factors for E- Assessment software

#UF	Proposed Factors	Sub Factors	Description
U1	Efficiency	----------	Efficiency indicates once users have learned the proposed prototype, how quickly they can perform tasks.
U2	Effectiveness	----------	Efficiency indicates once users have learned the proposed prototype, how quickly they can perform tasks.
U3	Functionality	Security	Security is the quality factor in dealing with those attributes of software that "bear on its ability to prevent unauthorized access, whether accidental or deliberate to program or data". Also, the most important quality factor to be taken into account when evaluating a CBAS proposed prototype.
U4	Satisfaction	----------	Satisfaction refers to the subjective responses from users about their feelings when using the software.
U5	Reliability of the Software	----------	It is important that no termination procedures should result in any loss of data. To ensure this, both student and proposed prototype files should be updated after each transaction, so that no data is lost if the test is terminated because of the machine or power failure.

Usability Criteria

To support Usability, there are various principles need to be followed in order, to make systems easy to learn and easy to use. These principles are defined by (e.g.Dix et al 1998, Nielsen 2003), More details are enclosed in Appendices A. These principles can be applied to the design of an interactive proposed prototype in order to promote its stability. Therefore, the purposes of adopting these principles are to give more assistance and knowledge to system developers (and the users) regarding the system design. Alongside the above principles, an important key additional factor is Utility. Utility refers to the functionality so users can *"do what they need or want to do"* Preece et al(2002, p. 16). In other words, *"does it do what users need?"* (Nielsen 2003). For that reason, Usability and utility are equally important in the development process and they need to be integrated"Nielsen(2003).

Usability/ UX goals

The peoples' interactions with interactive products must be optimized for enabling them to carry out their activities at work, school, and in their everyday life; Usability is generally regarded as ensuring that interactive products are easy to learn, effective to use, and enjoyable from the user's perspective Preece et al. (p.49, 2002).

a. Effective to use (effectiveness)
b. Efficient to use (efficiency)
c. Safe to use (safety)
d. Have good utility (utility)
e. Easy to learn (learn -ability)
f. Easy to remember how to use (memory-ability)

To describe it in more detail and provide a key question; for each goal. For example, *Effectiveness* is a very general goal and refers to how good a system is at doing what it is supposed to do. Usability goals are assessing how useful system is from its own perspective, while UX goals are concerned how users an interaction with a product from their perspective. The relation between the two in Figure (2-2) by Preece et al. (p.49, 2002) the Usability in the centre of the figure, the UX as a circle around the Usability. Any increase in Usability offset by an increase in the UX, the main goals of UX are [Satisfying - Enjoyable- Fun- Entertaining- Helpful- Motivating- Aesthetically- Pleasing-Supportive of creativity-Rewarding- Emotionally fulfilling].

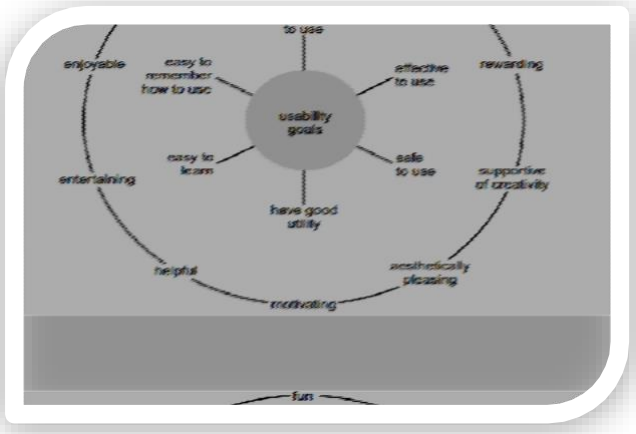


Figure 2-2: Explain the relationship between Usability and UX, UX goals shown in the outer circle, Usability goals are central by specific criteria.

2.3.2 User experience (UX)

Both broaden the concept of UX of:

- Include pleasure would be measured by the range of human responses.
- Include anticipated use and reflection on use would be measured by the circumstances.

Simply, about every interaction you have with what is in front of you, what if you feel at the moment of using it?, is *a "person's perceptions and responses resulting from the use and or anticipated use of a product, system or service"* Derom (2015).

The definition of Usability giving by Nielsen and Mack (1994, p.3) is: *" Usability is a fairly broad concept that basically refers to how easy it is for users to learn a system, how efficiently they can use it once they have learned it, and how pleasant it is to learn it, and how pleasant it is to use".* This definition furthermore relates Usability to user experience when it considers whether a system is pleasant to use, referring to the emotions or responses that users may experience.

2.4 Relevance of the previous research and this research

Mulvaney (2011)

Purposed the effect of CBAS on middle school age in learning students. He noted that the students toward to CBAS and positively interact with it, nevertheless, the tutors are divided into two teams: one of them agree to use it; while the others resist it. Finally, the study referred to "the CBAS will be an effective method to assess the students, and save tutors time in the future".

Simin and Heidari (2013:12732-12734)

Tried to extend the CBAS by pros and cons, their work provides a new assessment horizon for tutors and the development of assessment, it helps them to get a better understanding of using CBAS. This study emphasizes on CBAS replaced paper-pen assessment, i.e. "chalk and talk"; also, suggestions that the CBAS can provide a very good learning and the environment.

Tahir and Arif (2014)

Provided evaluation to design the user interface applications for children's education, and applied study on two different systems of children's applications are running Android. They attempted to re-examine the guidelines, menu interface, thus the development of the measurement model, the study pointed to the possibility of applying the standard on any system and other devices; also, offers a comprehensive structural model to evaluate the Usability.

Alqahtani (2015)

Investigated the effect on the Usability of multimedia metaphors, where she performed three experiments with 30 participants. Her work appeared the interaction of the user in terms of knowledge, satisfaction, and motivation; therefore, beneficial results for the development of interface get from the electronic assessment using multimedia metaphors, where the effectiveness of user interaction achieved with the interface evaluation. This work noted that in the education process needed to understand and use the technology in the field of evaluation.

Maqableh, et al. (2015)

Proposed a model depends on the acceptance of the technology model, and the theory of planned behaviour, which arrived at the model-based computer evaluated based on nine variables where they work a survey of 546 participants. They defined the factors that the student's attitude towards computer-based assessment, especially in higher education. At the end of the study, it emphasizes that the most likely computer-based assessment plays an important role in education if it is useful and easy to use.

Wilson (2015)

Presented an exploratory study, which it measures the user experience and Usability in University Libraries in the United Kingdom. This study has appeared that, the using theories of information environment as a method of building a user experience policy.

Kabir and Han (2016)

They presented that Point-of-sale proposed prototype because it provides fast and convenient ways of transactions for the business. These systems contain vital tasks such as online transactions, e-commerce facilities, security, taxes, various

management reports, and others. Thereby, it is important to ensure their software quality and grantee the effective usages of business functions.

They proposed an improved Usability evaluation model with a comprehensive view of Usability for a Point-of-sale system. They designed Usability scenarios for each factor and thus provided the corresponding questionnaires, and applied a case study of evaluating the Point-of-sale system in Bangladesh has demonstrated that the proposed model can provide a comprehensive evaluation of Point-of-sale system from 12 Usability factors. Also, different demands from a different type of customers are being revealed by the model. They proved its effectiveness through their case study, also they highlight different software Usability depends on different types of users. They cited to would like to apply the similar methodology to build enhanced models for evaluating other software quality attributes

Rajanen and Clemmensen (2017)

They examined the views of UX professionals on the definitions of Usability and UX and compared the findings between five countries and within different socio-cultural groups. Make the distinction between Usability and UX clearer and firmer, and provide guidelines on using the two concepts in design and evaluation within organizations. They also contribute by pointing out that UX professionals refer to a variety of characteristics and attributes associated with Usability and UX that parallels the struggles in HCI research on finding the best ways to capture the essence of these concepts, as they evolve in time.

As a result, Usability characteristics and guidelines are important for interface design of E-Assessments application Tahir and Arif(2014). Also, Usability is a significant factor evaluation website and applications Rajanen and Clemmensen (2017) Therefore, A Measurement Model based on Usability Matrices (MMB-UM) for measuring the Usability of CBAs applications will be introduced in Chapter 3 and Chapter 4.

2.5 Conclusion

It is obvious that the previous studies of researchers have focused on the HCI fields is separated from the E-Assessments fields. The assessment is an essential process and important in the field of education has many special applications and tools. Briefly, this research differs from other research in combining between two areas: the

Electronic-Assessments (E-Assessments), and HCI perspectives, therefore, addresses the overlap between the Usability and UX. It will introduce a new model to evaluate the proposed prototype of CBAs, where the measurement model based on Usability metrics (MMB-UM).

Chapter3

The Proposed Design of CBAs and MMB-UM

3.0 Introduction

It is important for the researcher to determine the methodology that will be followed, based on Plantak and et al(2010) cited that: Kothari (1990, p. 8) "The research methodology is a way to systematically solve the research problem" and "research methods do constitute a part of the research methodology". Through the previous literature review and relevance work, appeared clearly the methodology that will be used in this research,which is divided into three parts: design of proposed prototype Computer-Based Assessments (CBAs), proposed Measurement Model Based on Usability Metrics (MMB-UM), and utilize MMB-UM for evaluating Human-Computer Interaction (HCI) perspectives : Usability and User Experience (UX). This chapter explains the method and techniques, which used in this research.

3.1 Participants in the Experiments

A random sample of students at the Information Technology Faculty from various specializations, semesters **(III, IV, V)**, males and females, based on recommendation number of participants, which proposed by Ginny(1999) that cited in www.usability.com(2016), Table (3-1) explains recommended number of participant , which required to make reliable estimates to cover the Usability problems Tahir and Arif(2014).

Table 3-1: Recommended Number of Participants

Source: Testing Technique based on Ginny Redish on the site techsmith.com

Test type	**Summative Testing**
How many	**6_12** Users
Metrics and measures	More formal metrics based on Usability goals/MMB-UM
Why	Measure success of new proposed prototype design
When	At end of proposed prototype
How often	Once

3.2 Instruments of Design

Any experiment should have suitability tools for obtaining exact results, for this reason, the experiments in this research require the following tools to implement successfully.

3.2.1 Design of proposed prototype CBAs

There are many tools and applications available for the design of the (CBAs), Free tools such as : Course Lab, hot potatoes, etc.., and trade tools such as quest base, etc. , also; free sites provide the assessment service; online tests. In this research, chose the Course Lab tool to accomplish CBAs, which is resembles in the environment of HCI, where it has the objects, modules, and interfaces, interactive and in addition, that he has a wonderful and nice user interface. Figure (3-1) illustrates the course lab tool page the right sees interactive objects and to the left modular *"Course Lab is powerful, yet easy-to-use, e-Learning authoring tool that offers a programming-free WYSIWYG environment for creating high-quality interactive e-Learning content that can be published on the Internet, Learning Management Proposed prototype (LMS), CD-ROMs and other devices"* (www.courselab.net).

Figure 3-1: Home Page of Course Lab Tool to create CBAS

3.2.2 Testmoz Site for Online Assessment

Creating E- Assessments requires high cost, workshops and tools, websites, and program, etc… Inasmuch of the limited possibilities in the University, especially in the present time, for these reasons were selected Testmoz the site (www. Testmoz. Com), pointed out to that*: "Testmoz is a test generator that supports four question types, automatic grading, a really simple interface, and detailed reports. It is free and does not require you (or your students) to register*." Figure (3-2) explains the home page of Testmoz site.

Figure 3-2: Home Page of Testmoz to create online assessment

3.2. 3 The Measurement Model Based on Usability Metrics (MMB-UM)

There are many famous measurements, models to measure both factors (Usability, User Experience (UX)) in this dissertation the MMB-UM chosen to measure (Usability, UX) and the reason for this selection was the ability to measure the two factors together, reverse all famous measurements that measure each factor separately.

In addition, the MMB-UM selected for the following reasons:

- It considers the first time to use in the Electronic- Assessments (E-Assessments)
- User satisfaction in the measurement is equal to the UX.
- It presented the Usability as objective, and the UX as subjective.

In the Figure (3-3) explains the MMB-UM model for measuring the Usability of Electronic Assessments (E-Assessments). The proposed model consists mainly of three phases: The first phase explores the literature review as clarified in *Section (2.3.1.2)* and presents the Usability characteristics and guidelines for E-Assessments. In the second phase Goals, Questions, and Metrics of Usability. In the last phase, the metrics are separate into the objective (Usability) and subjective (UX) metrics, which used to develop two measurements instrument task list and questionnaire respectively.

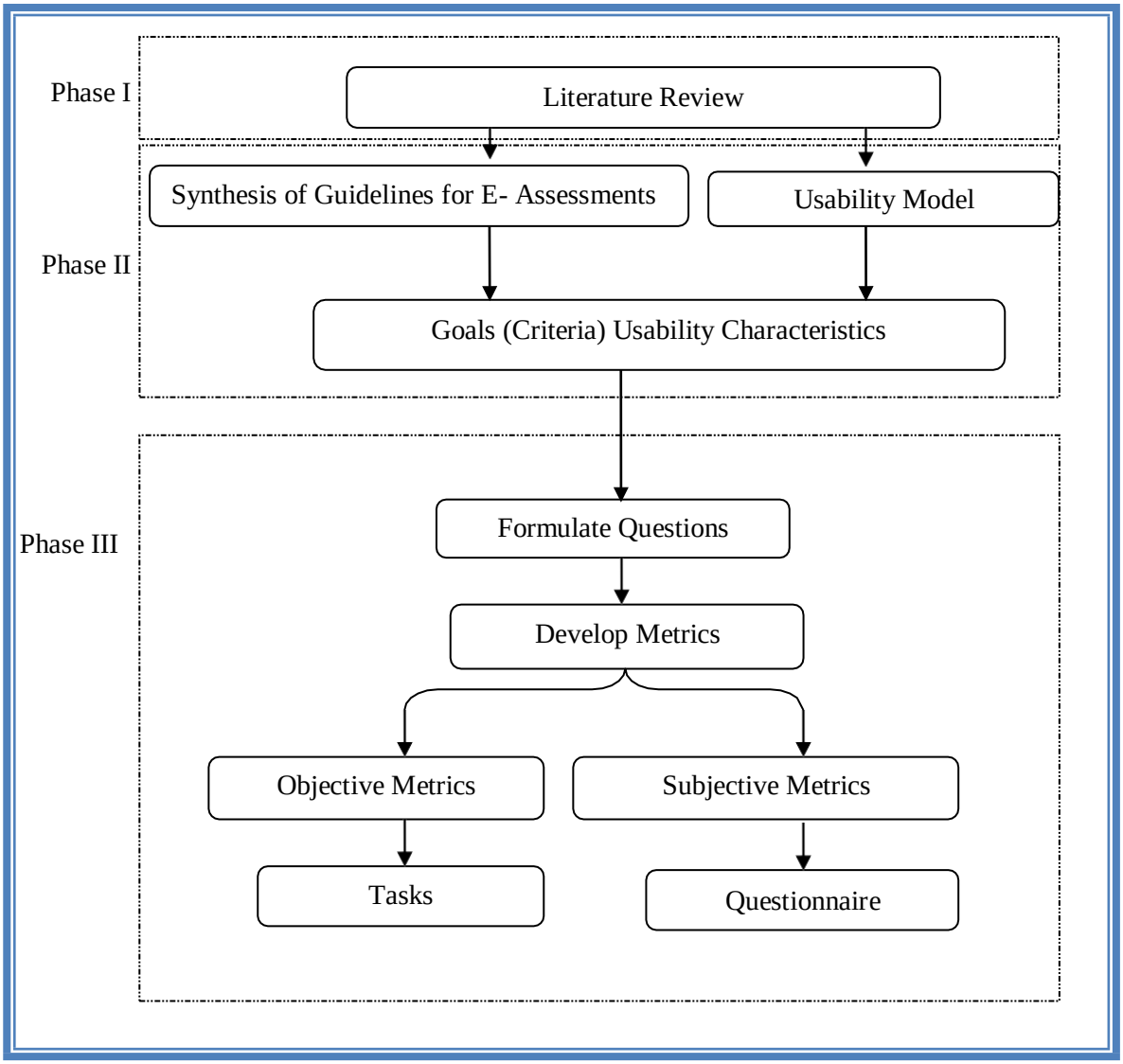


Figure 3-3: General Structure of Measurement Model based on Usability Metrics for proposed prototype of E-Assessments

Phase II consists goals and guidelines were being to formulate by this research, in Table (3-2) shown the resulted goals, questions, and metrics of Usability for E-Assessments.

Table 3-2: Phase II is illustrating Usability Characteristics, Goals, Questions and Metrics

Usability Characteristics	Goals	Questions	Metrics
Effectiveness	Interactivity	Is it easy to interact with the proposed prototype?	Number of mistakes during interaction
Understandability	Input/output	Does the proposed prototype provide easy ways of input for students?	The number of mistakes to enter
Efficiency	Time required	How much time taken by the application to load?	Time is taken to load application rating scale for time response
Satisfaction	Attractive	Is the proposed prototype attractive for students?	Rating scale for Attractive screen design
Security	Secure	Is the proposed prototype safety the confidence of tutors?	Number of mistakes in secure

Phase III includes Measurement Instruments (*"The subjective and objective metrics from the previous phase are used to develop two evaluation instruments user satisfaction questionnaire and task list respectively"* Tahir and Arif(2014). As shown in Table (3-3) and Table (3-4). The questionnaire has 26 questions appropriate for E-Assessments, based on the Usability factors, which defined in Table (3-2).

Table 3-3 (a): Phase III Questionnaire of MMB-UM for CBAS

Question Index	Question
1	I found it easy to understand this application.
2	The application is too slow I had to wait for a response to continue
3	The application took a lot of time for loading

Table 3-3 (b): Phase III Questionnaire of MMB_UM for CBAS

Question Index	Question
4	The application provides a visual display to show the loading process.
5	The application gives feedback on whether my answer is correct or wrong
6	The application does not provide appropriate feedback to my answers.
7	I was comfortable with the screen orientation of the application.
8	The main menu of the application is confusing.
9	The app provides clear and understandable navigation keys such as back/next buttons to move to the previous / next screen
10	The application provides useful help information
11	It was difficult to find help
12	It was difficult to understand the language used in the application
13	The topics/concept and information was understandable
14	I need to remember a lot of information throughout several actions to perform a task.
15	It is easy to complete the tasks without much effort
16	It is difficult to learn to use the application
17	The application provides a progress report/result of my performance in every activity
18	The application gives error messages that clearly tell me how to fix problems
19	It was easy to read the text in this application
20	The text size used in this application is too small
21	It is easy to find the information I needed

Table 3-3 (c): Phase III Questionnaire of MMB_UM for CBAS

Question Index	Question
22	I find the design of application attractive
23	The colours used in this application are not attractive
24	The icons and buttons used are attractive and recognizable
25	The application gives interesting rewards on my performance
26	Overall, I enjoyed using the application.

Table 3-4(a): Phase III Task List of MMB-UM for CBAS

Task Index	Task
T1	**Check for interactivity** a. Check for user interaction with the application b. Check on availability of communication tools c. Check of usage of gestures
T2	**Navigation activity** a. Check of main menu presence b. Check for scrolling c. Check for hierarchal menu d. Check for navigation keys
T3	**Check for the time** a. Loading application b. Task

Table 3-4(b): Phase III Task List of MMB-UM for CBAS

Task Index	Task
T4	**Check for adequacy of Help** a. Task-related clues b. Tutorials c. Help icon
T5	**Check for cognitive load** a. Identify a link or icon usage b. Check for suitability of language c. Check for suitability of content
T6	**Check for learning potential** a. Check for presence of alternative learning options b. Check of assessment/result availability
T7	**Check for personalization/customization** a. Check for availability of settings option
T8	**Check suitability for reading**

3.2.4 Design of Questionnaire to MMB-UM

Data compiled and analysed according to answers to the questionnaire; it is given to twenty students of Information Technology (IT). Qualitative data collected through a questionnaire on the internet by Google as shown the icon in Figure (3-4), and utilized Microsoft Excel 2007 to diffract these data. The questionnaire is part of the Usability Measurement Model usability Metrics (MMB-UM) has selected 26 questions from the commensurate with the CBAs, distributed the link of the questionnaire to the students, so each student can access it from any place and answer based on his experience. The questions will be explained and the results clarification in detail in chapter 4.

Figure 3-4: Google Forms to create online Questionnaire

3.3 Procedures of proposed prototype CBAs Design

At the beginning of the experiment, take into account several points have been deduced from the Usability Testing Basic method:

1. Select the background: description of the proposed prototype, also the feedback request. Where the CBAs is considered a newly proposed prototype at the University of Benghazi, need to know the extent to accept students and tutors, and its impact on the educational process.
2. The participants: the required qualities of the participants, and characteristics. "*The selection of a random sample of students who were eager at experience".*
3. Usability objectives: identify the factors that I need to make sure that they are effective to support the proposed prototype in the future."Competency, security, trust, and learning... etc. are the most important factors for CBAs".
    - Keywords: specific questions we need participants to answer.
    - Timeline: special when the proposed prototype is ready for evaluation.
    - Additional information: user reviews that we need to consider in the future, ensuring proposed prototype quality and effectiveness.

Briefly, the procedure, which followed in this experiment as:

- Done an agreement with the tutors of the experiments courses, to test the students in the multiple choice questions **(MCQs)** summative assessments CBAs; also, online assessments. Collecting questioner online, it; the measurement was chosen for this research.
- Conducted the experience of CBAs for students in the specific environment (lab); downloaded the executive file on twenty computers, compared the CBAs with PBAs and online assessment.
- In the absence of complete facilities, connectivity between devices in the lab for online assessments; did that, giving the link of assessments to the students.

3.4 Limitations of the proposed prototype CBAs Design

It is important to note that this research does not address the below challenges which determined by the researchers (e.g. Latour and Sarre,2018) field of designing CBAs: Requires high intuitive, design interface, easy-to-learn, and the interface should not distract students from test content; and navigate easily; support for a wide range of test-taking strategies as an example the essay exam; the students in various semesters and the courses need to a different requirements for design their exam accommodating the students with disabilities.as a result, this research could not cover all these challenges.

The experiments were limited; because the participants were few, so they didn't reflect the views of all students.

3.5 Conclusion

This chapter provides an overview of the methodology used in the experiments, which will be presented in the next chapter to achieve the objectives of this research, and the tools used to design the proposed model of CBAs, the reasons for its selection, and the challenges in the E-Assessments researchers are facing.

Chapter4

Experimental Results

4.0 Introduction

Provide important information about the experiments that were conducted during this research, it will introduce the methodology which referred at chapter 3 in order to reach the objectives of this research, then the results of the experiments which implemented through this research.

4.1 The experiments aims and objectives

It is important to encourage the students and the tutors to interact with an Electronic Assessments (E-Assessments); also, an emphasis on the E- Assessments are a suitable tool to enhance education. Moreover, the aims of experiments were conforming to the overall aims of the research. This experiment aims to achieve the following:

1. Examine the effect of electronic assessments on the education process.
2. Investigate if electronic assessments are abolishing the phenomenon of cheating or not.
3. Examine if Measurement Model Based on Usability Metrics (MMB-UM) suitable for electronic assessments.

Usability main factors, which evaluate are (efficiency, user satisfaction equal UX, effectiveness) by:

- Measuring efficiency (time spent by students to complete the required tasks);
- Measuring effectiveness (by calculating students correctly performed tasks);
- Measuring the users' satisfactions;

- Measuring UX used an A Measurement Model Based on Usability Metrics (MMB-UM).

4.2 Tasks of the Experiments

1. Formulate experiments hypotheses.
2. Measure the efficiency, in terms of time, which taken by each student to complete his exam; also, the time of the tutors to correct students' answers.
3. Measure the effectiveness of E-assessments on students‘ answers.
4. Measure the user satisfactions, for all experiments by allowing students to see feedback on their answers, and comfortable the interfaces; in addition, the tutors might be confident at E-Assessments.
5. Analysis the results to compare between Paper -Based Assessments (PBAs) and Computer -Based Assessments (CBAs).

4.3 Hypotheses

Several Usability measurements have been improved to evaluate the quality of E-Assessments, whence saves time, low cost and effort. Based on the factors of Usability presented in this research, it is expected that the Measurement Model Based on Usability Metrics (MMB-UM) will provide a satisfaction and confidence for the tutors especially in term of security. Consequently, the hypotheses were formulated regarding of successful E- Assessment:

- **Ease of interact:** how the CBAs and online assessment are enhanced for students and tutors.
- **The efficiency of use:** how quickly students can locate the answer through the use of (CBAs and online assessments) for the first time.
- **Effective test:** how effective the use of E-Assessment in the education process, to evaluate students.
- **Satisfaction:** the student will be more satisfied with the E-Assessment is more usable if he enjoys when achieving a given task.

- **Security:** tutors need to be sure that electronic assessments ensure privacy and high confidentiality**.**
- **Portability:** how can install and transferable of E-Assessments.

The hypotheses were formulated regarding the aspects of the success of E-Assessment in higher education:

H1: Electronic assessments will be accepted in higher education by tutors and students.

H2: Measurement Model Based on Usability Metrics will be effective to measure CBA.

H3: Measurement Model Based on Usability Metrics will be effective to measure online assessments.

4.4 Design of Experimental Condition

To achieve the suggested hypotheses, electronic assessments have been designed and advanced according to guidelines. The guidelines pertaining to design CBAs (almaarik.wordpress.com) , and online assessments (elearningksu.wordpress.com). The electronic assessments provided two different versions: an assessment CBAs offline, and online assessment.

Both versions of the empirical were designed to provide similar information about the exam. This was represented in the form of two types of questions: true or false questions and multiple choice questions.

4.4.1 Computer Based Assessments (CBAs)Design

CBAs is designed the colours white and blue screen designed, comfortable to student's eyes, and every question has time only 10 seconds. The first page displays the title of the exam and instructions of the exam entry button; all that illustrated in Figure(4-1), the second pages consist of three labels Figure (4-2) shows that:

- To chose the department;
- To enter the name;
- To enter the number;

With "click button" to access the exam. Every page's design has incorporated the feature by dividing the screen, question by question, in order to avoid overlapping of questions, so the user can easily select answers on the screen. There is timeline 10 second to answer each question when the student answer submits button appears to store answer; stop timeline, and student transfer to the next question, all that is illustrated in Figure (4-1) text of the question in a font size 18 for the test.

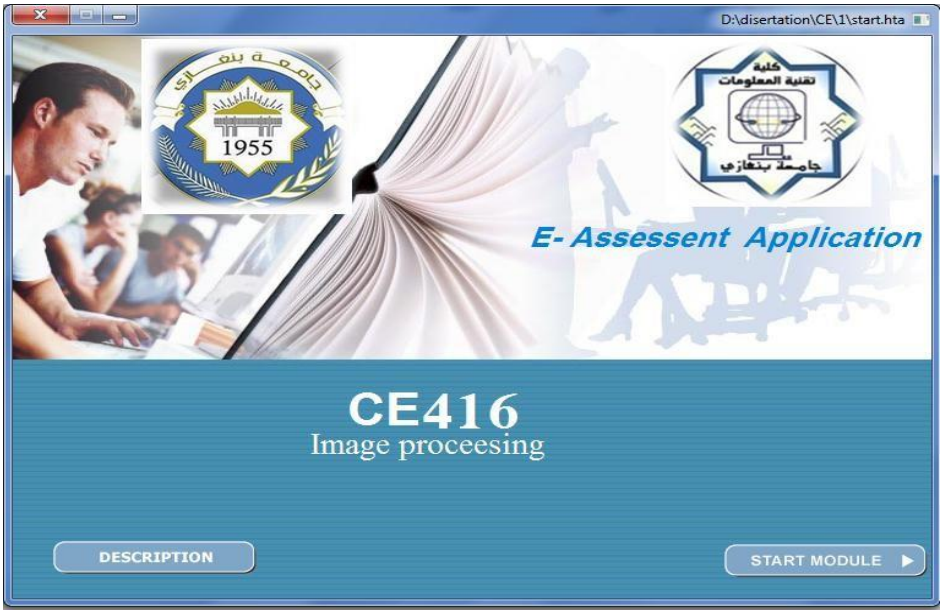


Figure 4-1: Illustrate the first page of CBA offline

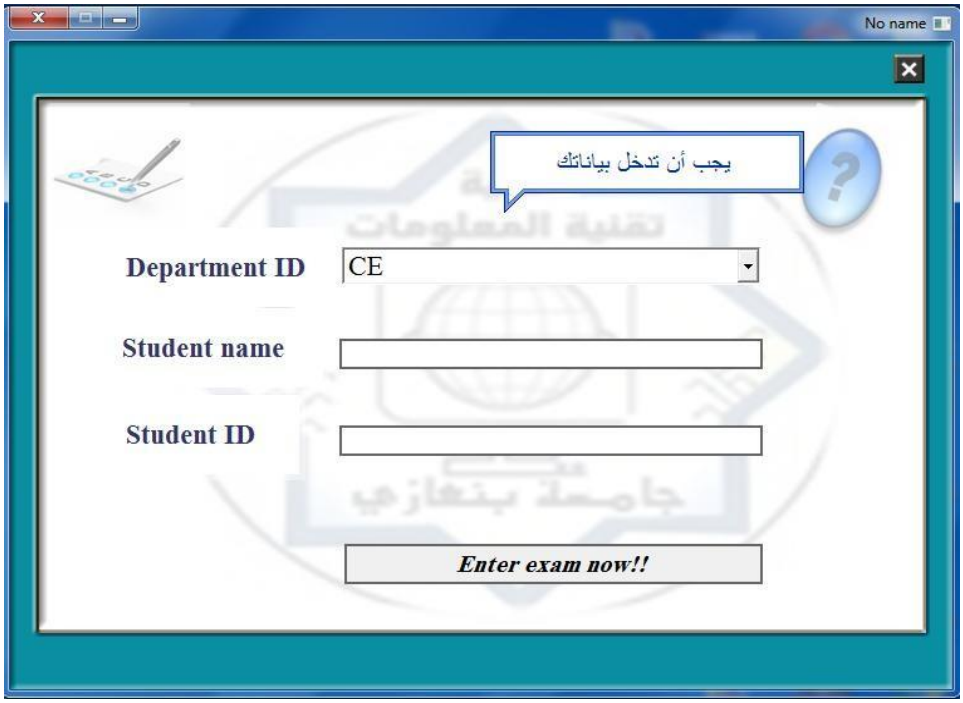


Figure 4-2: Illustrate the second page of CBA offline

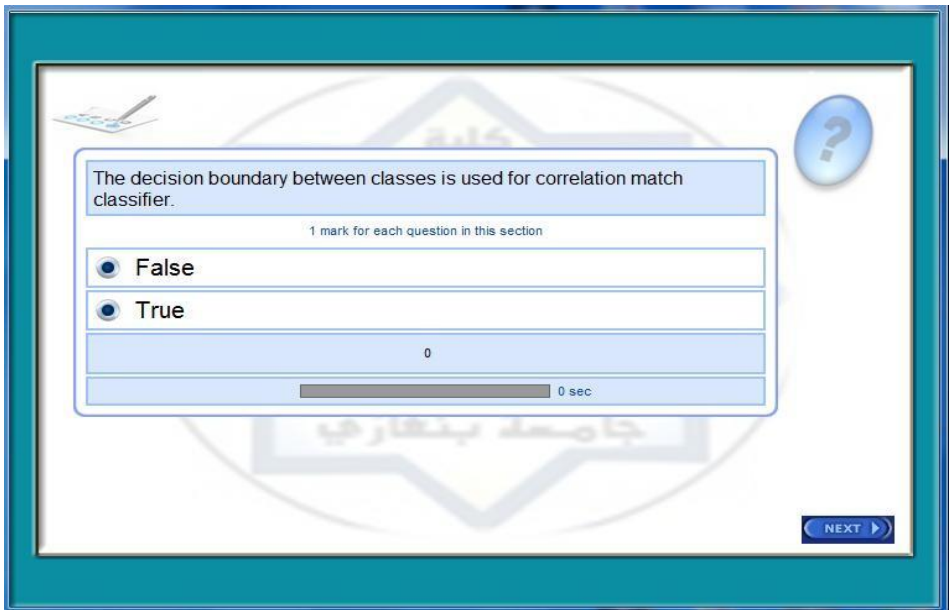
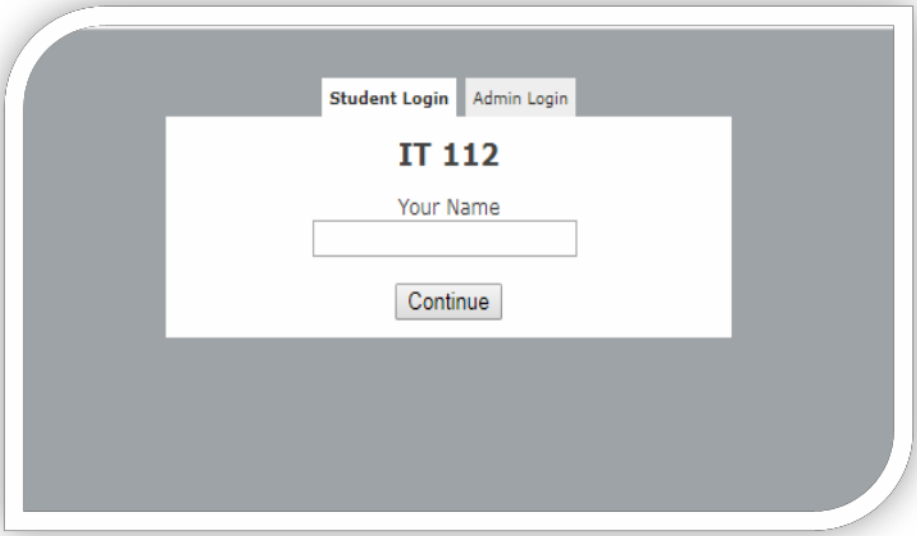


Figure 4-3: Illustrate page question, and time line

4.4.2 Introduce of Online Assessments

The site offers free E- Assessments services. It contains two modules, one for the students as the name of the course to take the exam, a box for entering the student's name and number, and the second-page access button. All this is shown in Figure (4-3).

Move to a page divided into two parts; the instruction section at the top and all questions are displayed at the bottom. As shown Figure (4-4).



,

Figure 4-4: Students page of online assessments

Figure 4-5: Illustrate questions page, up the instruction of the exam

Another module for tutors where the PIN is entered, that takes them to a page that contains a set of links, including a link to the students' grades, the hours they enter the exam, and the time which takes for each student to answer the exam. Statistics

for the correct answers and the wrong answers for each student. This is shown in Figure(4-5), Figure (4-6), and Figure (4-7) below.

Figure 4-6 Tutor login

Figure 4-7: Links for the tutors

Figure 4-8: Statistics of answers for each student

4.5 Variables of the Experiments

In the experiment, design identified the variables into three types: independent, dependent and controlled, as follows:

4.5.1 Independent variables

The values that can be changed or controlled. It provides the "input" which is modified to change the "output"(www. yourdictionary.com). The independent variables are important to represent aspects of the experiments, which are expected to be the reason for the results. The independent variables in these experiments were the tutors and the students; also, they had participated by their feedback at the end of the experiments.

4.5.2 Dependent variables

"The values that result from the independentvariables"(www.yourdictionary.com)

The dependent variable in this experiment was used to measure effort and efficiency,by the time of question correcting ; effectiveness by recording the time

spent to complete the exam by each student; UX satisfaction using a questionnaire based on MMB-UM and analysis the responses by using the Likert scale.

Table 4-1: Dependent variables

Variables	Measure
Time of question correcting	Efficiency, Effort
The time spent to complete the exam	Effectiveness and students' performance
User Satisfaction	Satisfactions

4.5.3 Controlled Variables

"The variables expected to influence the experiment process were controlled". The variables in this experiment were:

- **Tasks:** every student has information about tasks that will request from him in the experiments.
- **Courses:** the subjects tested in the experiments were similar to interface versions with questions in the same level, except the orders of questions are different.
- **The consciousness of tasks:** students were conscious of the tasks that would be provided to them.
- **Time:** students had sufficient time to complete all particular tasks Consequently, a task completed within the time allocated would be regarded as successful or else the task would be conceived as unsuccessful.
- **Condition familiarity:** each student was given the same instructions, and it was the first time for him to do the exam.
- **Consistency:** The execution of the experiment interfaces was examined among the similar students on a personal basis to all students.

4.6 Experiment Sessions

The experiment was carried out to ensure the MMB-UM model is reliable and effective for evaluating the Usability and UX of E-Assessments. The MMB-UM model was conducted to test (User satisfaction questionnaire and task list) Therefore, both objective and subjective metrics were employed in these experiments. To validate the model used two type of E-Assessments (CBAs and online) in the experiments two tasks were executed for each assessment by each student

4.6.1 Experiment 1 (CBAs)

The experiments were conducted in a quiet and comfortable lab at Benghazi University. Before the experiments began the lab was prepared (equipment, tools,.Etc.), and installed the CBAs on computers. The exam took between 15 and 20 minutes with a mean time of 17.5 minutes, i.e. the exam time was distributed as follows:

Student participants listened to the Instructions around 2 minutes, then they started to answer the exam in 10 minutes. Upon completion, students filled the questionnaire which took about 6 minutes. The CBAs experiments were included 20 students, to take the exams in the following courses (CE416, CN281, IS361, IT112), thereafter required from the students interact with the proposed prototype of CBAs in terms of Usability, and background colours.

The second evaluation of the proposed prototype of CBAs through a questionnaire contains 26 questions to judge the user experiences. During Usability test each student was asked about the proposed prototype of CBAs, that was protocol *"Concurrent Think Aloud"* to know their compatibility with the proposed prototype and took their observations.

4.6.2 Experiment 2 (Online Assessments)

From their homes, the students answered the exam questions online that built confidence among the students and the tutors, the students had given a link to the

exams on the site Testmoz for the same courses they had done in the proposed prototype of CBA, and gave them the following tasks:

The First Task: Answered questions.

- Please open program www.Testmoz.com.
- Enter your name and number in the box.
- Enter login.

The Second Task: Filled a questionnaire.

4.6.3 Design the Scenario for Security, Usability Factor in E-Assessments

"*Scenario is an effective means of capturing the software quality attributes*" Kabir and Han (2016). Might design scenario for each factor of Usability which had clarified in Chapter 2 section 2.3.1.2, for example, security is an important factor in design E-Assessments, for this reason, were selected security factor scenario design, which described Table (4-2).

Table 4-2: Usability Scenario for security factor

Portion of Scenario	Possible Values
Source	*Students/Tutors*
Catalyst	*Try to access the site by using the URL and entrer student information*
Environment	*Runtime E- Assessments online*
Artifact	*Testmoz site*
Response	*Details of exams are retrieved from the database depending on its code*
Response Measure	*Security (Students are accessed by URL, they do their exam, submit their answers, their results transmit to tutors account unravel time access for each student)*

Its scenario diagram is shown in Figure (4-9) shows a usability factor's scenario of security for E- Assessments. In the scenario, the students could access to the Testmoz site by using the URL, and students enter their information, they take the

exam copy from a database and automatically transmit their answers to tutor account.

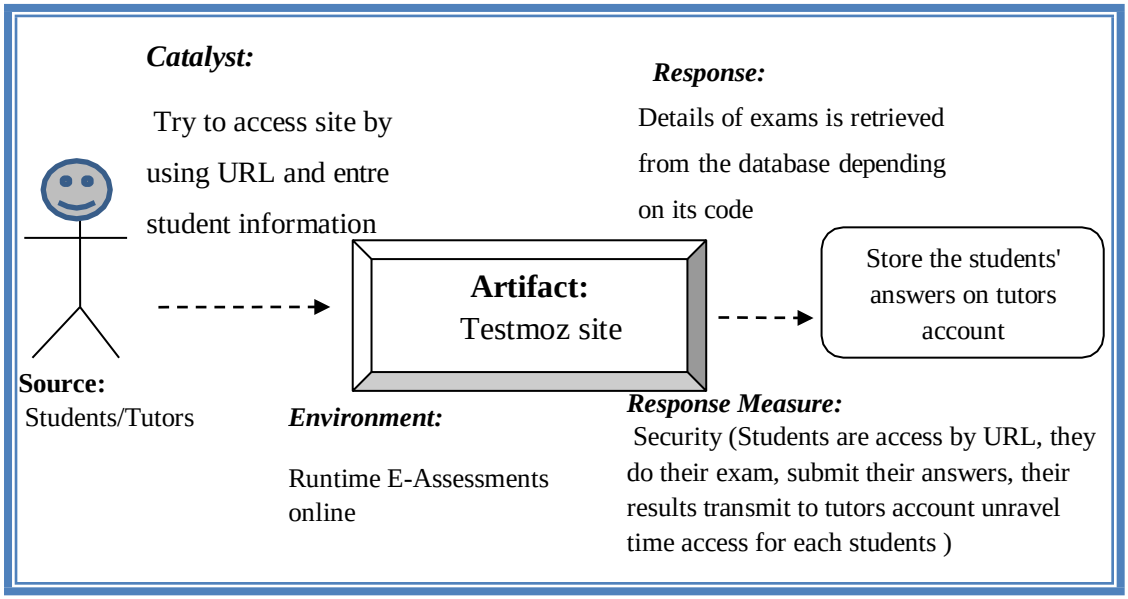


Figure 4-9 Usability Scenario for security factor

4.7 The MMB-UM Results

These experiments were carried out to ensure that the MMB-UM model is reliable and effective for evaluating the Usability and UX of E-Assessments by(User satisfaction questionnaire and task list) which are sections of the model.

To validate the model this experiment used two types of assessments CBAs offline and online assessments that designed for higher education. Usability evaluation the subjective data were collected using the five points from the Likert scale (that illustrate in Table (4-3)) user satisfaction questionnaire developed in the MMB-UM model and objective data were collected through the task list to prepare tasks for each assessment. Obtained the results for both subjective and objective metrics are presented separately and comparison of results presented for both assessments to check significant differences in Usability and UX of the two E-Assessments.

Table 4-3: Satisfaction Level and values

Satisfaction Level	Level's Value
Excellent	*5*
Good	*4*
Neutral	*3*
Fair	*2*
Poor	*1*

4.7.1 Objective Usability Results of Experiments

In Tables below, labels O1-O6 were used to represent the objective metrics. For a comparative analysis of the two E-Assessments, the results are presented in Table (4-4) to Table (4-7), which calculated by Usability metric equations , www.usabilty.com (2015) had indicated to (ISO/IEC 9126-4) that recommends that Usability metrics should include obtaining outcomes effectiveness, efficiency, and user satisfactions.for more details enclosed appendix c.

The data for objective measures were collected during the Usability measure and summarized the data for each of the six objective metrics. The mean score for each measure is presented in Table (4-4) until Table (4-7) for both E- Assessments for four courses.

Table 4-4: Results for Objective Metrics for CE416 Course

Objective Code	Objective Metrics	CBAs Mean	Online Mean
O1	Number of mistakes during interaction	1.1	2.4
O2	The number of mistakes to enter	0.00	0.02
O3	Time is taken to load the application	0.031	1.037
O4	Rating scale for time response	0.20	0.57
O5	Rating scale for the attractive screen design	0.8	0.9
O6	Number of mistakes insecure	0.01	0.7

Table 4-5: Results for Objective Metrics for CN281 Course

Objective Code	Objective Metrics	CBAs Mean	Online Mean
O1	Number of mistakes during interaction	0.15	1.4
O2	The number of mistakes to enter	0.01	0.06
O3	Time is taken to load the application	0.023	1.12
O4	Rating scale for time response	0.30	0.67
O5	Rating scale for the attractive screen design	0.5	0.6
O6	Number of mistakes insecure	0.03	0.9

Table 4-6: Results for Objective Metrics for IS361 Course

Objective Code	Objective Metrics	CBAs Mean	Online Mean
O1	Number of mistakes during interaction	1.1	1.6
O2	The number of mistakes to enter	0.02	0.92
O3	Time is taken to load application	0.006	2.01
O4	Rating scale for time response	0.2	0.5
O5	Rating scale for attractive screen design	0.8	0.6
O6	Number of mistakes insecure	0.02	1.9

Table 4-7: Results for Objective Metrics for IT112 Course

Objective Code	Objective Metrics	CBAs Mean	Online Mean
O1	Number of mistakes during interaction	2	2.4
O2	The number of mistakes to enter	0.00	0.02
O3	Time is taken to load the application	0.45	2. 69
O4	Rating scale for time response	0.40	0.63
O5	Rating scale for the attractive screen design	0.9	0.7
O6	Number of mistakes insecure	0.6	4

Summary

The comparative analysis is carried out to determine Which is the best type of E-Assessments in the experiments of whence Usability. The results indicate that online assessments have higher failure and a number of mistakes of all objective measures shown in Figure (4-9)until (Figure (4-12), except for navigation.

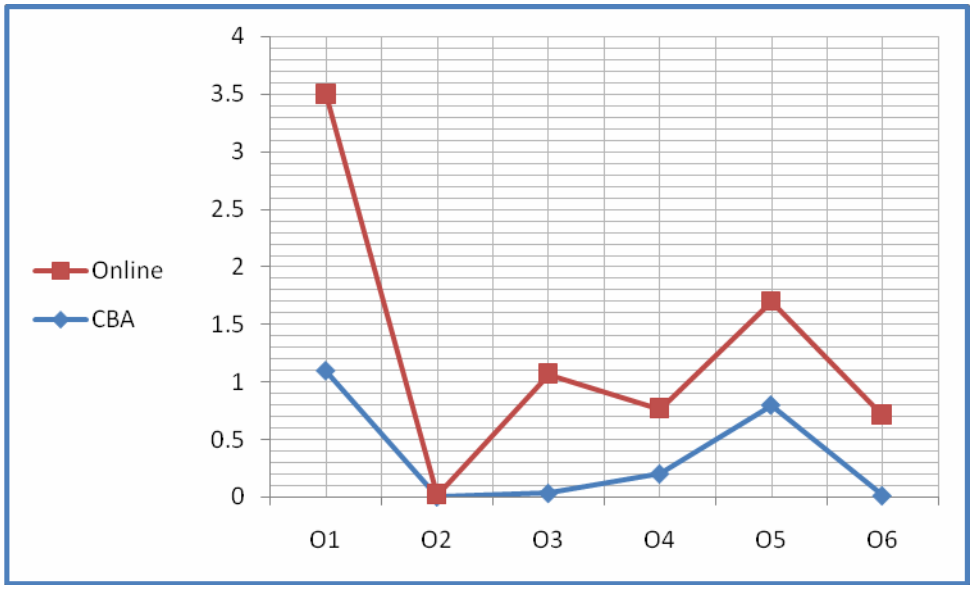


Figure 4-10: Objective Usability of E- Assessments for CE416

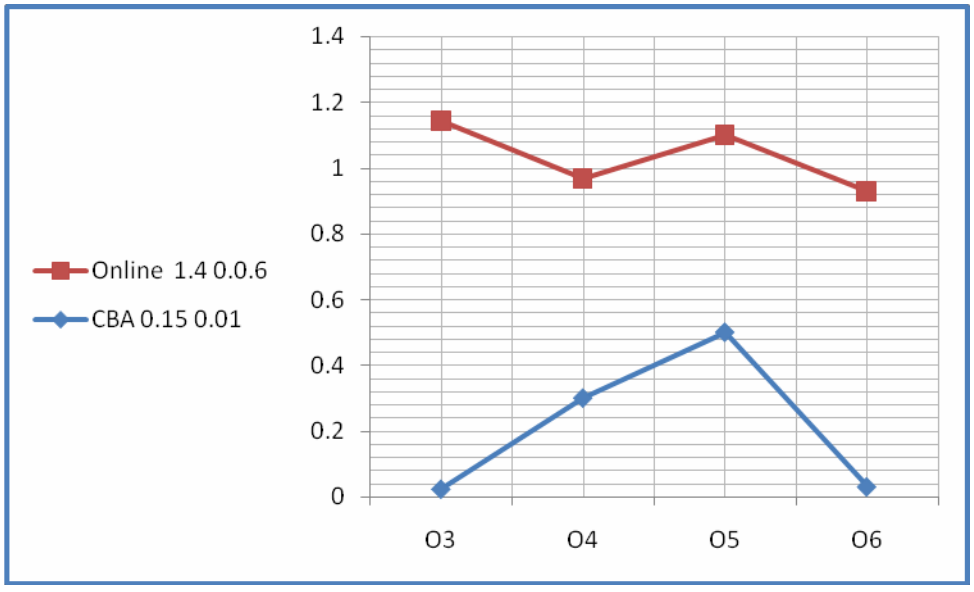


Figure 4-11: Objective Usability of E- Aassessments for CN281

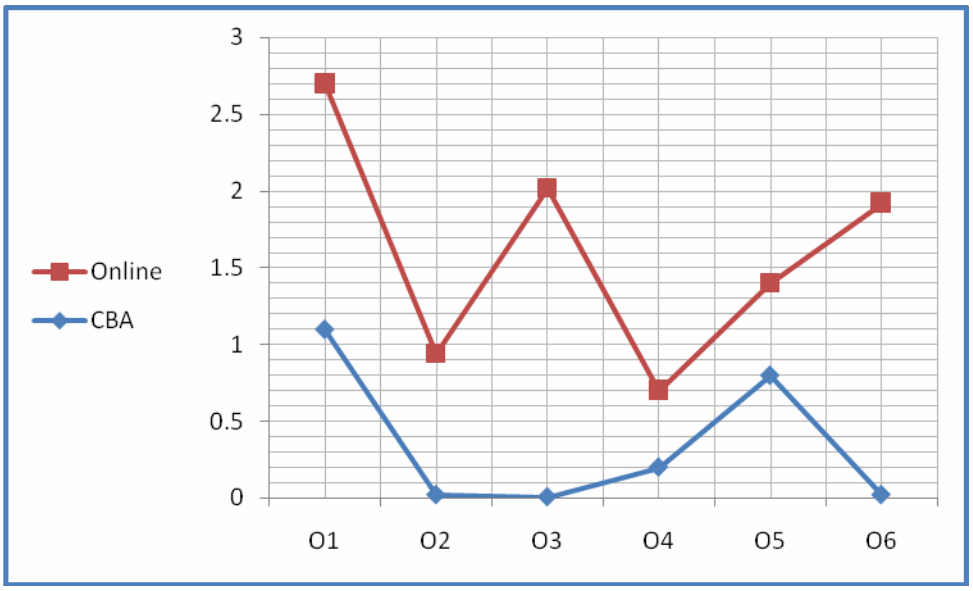


Figure 4-12: Objective Usability of E- Assessments for IS361

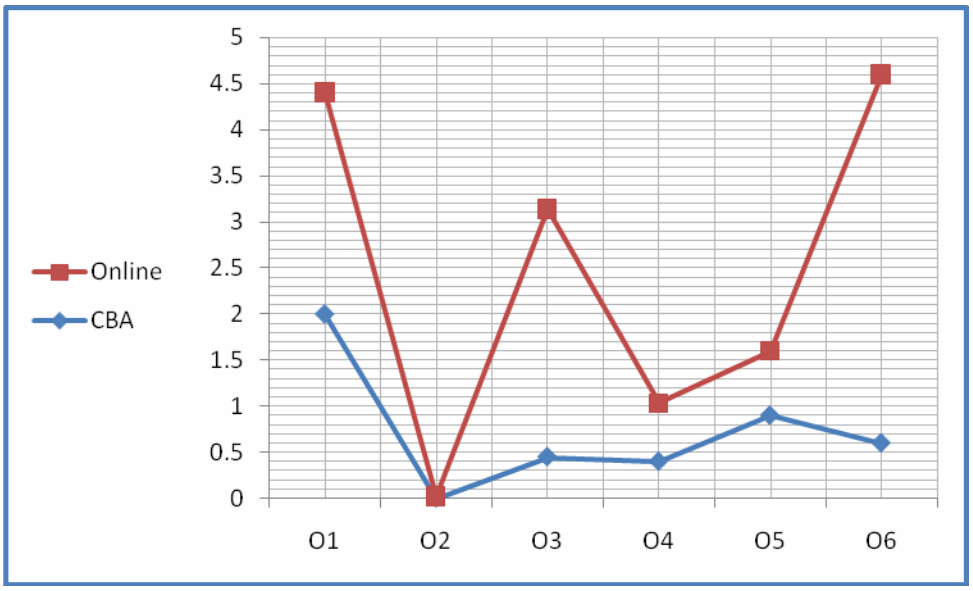


Figure 4-13: Objective Usability of E- Assessments for IT112

4.7.2 Subjective Usability Results of Experiments

The data for subjective measures were collected through five points Likert scale user satisfaction questionnaire presented in the model. The questionnaire was filled by the students after performing tasks, at the end of the test session for each assessment. The analysis of data based on a Likert scale (for more information, recur to Table (4-3) shown Table (4-8)to Table(4-11) satisfaction questionnaire was done according to the statistical procedure described by Boone et al(2012). For each subjective metric, the questions from satisfaction questionnaire were matched with the metric (for example, from Table (4-8) to Table (4-11) question number 25 relates to subjective measure *"Ease of readability"* and mean score for these questions was recorded for each of these subjective metrics. A higher score indicates greater satisfaction level for users. In Table (4-8) to Table (4-10) presents the results of subjective measures for each course. The labels "S1 to S23" are used to represent each subjective metric.

Table 4-8 (a): Results for Subjective Metrics CE416

Subjective Code	Subjective Metrics	Question No	CBA Mean	Online Mean
S1	Rating scale for multimedia usage	27,28	33% 1.65	22% 1.1
S2	Rating scale for appropriate feedback	8,5	42.9% 2.145	60% 3
S3	Rating scale for navigation	11	16.7% 0.835	33% 1.65
S4	Rating scale for the main menu	10	50% 2.5	60% 3
S5	Rating scale for pedagogic feedback	6	70% 3.5	80% 4
S6	Rating scale for easy to understand output	1	30% 1.5	43% 2.15

Table 4-8 (b): Results for Subjective Metrics CE416

Subjective Code	Subjective Metrics	Question No	CBA Mean	Online Mean
S6	Rating scale for easy to understand output	1	30% 1.5	43% 2.15
S7	Rating scale for loading application	4	77% 3.85	62% 3.1
S8	Rating scale for time to respond	3	66% 3.3	21% 1.05
S9	Rating scale for task effort	17,18	50% 2.5	55% 2.75
S10	Rating scale for finding help	13	14.3% 0.715	18% 0.9
S11	Rating scale for appropriate language	15	77% 3.85	88% 4.4
S12	Rating scale for appropriate content	16	39% 1.95	46% 2.3
S13	Rating scale for ease of learning	19	52% 2.6	66% 3.3
S14	Rating scale for suitability for all users	21	33% 1.65	22% 1.1
S15	Rating scale for performance assessment	23	49% 2.45	50% 2.5
S16	Rating scale for error messages	24	16% 0.8	19% 0.95
S17	Ease of readability	25	22% 1.1	35% 1.75
S18	Satisfaction with text	26	80% 4	74% 3.7

Table 4-8 (c): Results for Subjective Metrics CE416

Subjective Code	Subjective Metrics	Question No	CBA Mean	Online Mean
S19	Rating scale for engagement	34,35	50% 2.5	63% 3.15
S20	Rating scale for screen layout	29,30	49% 2.45	57% 2.85
S21	Rating scale for attractive screen design	31	33% 1.65	19% 0.95
S22	Rating scale for interface color	32	17.2% 0.86	20% 1
S23	Rating scale for icons and buttons	33	22% 1.1	30% 1.5

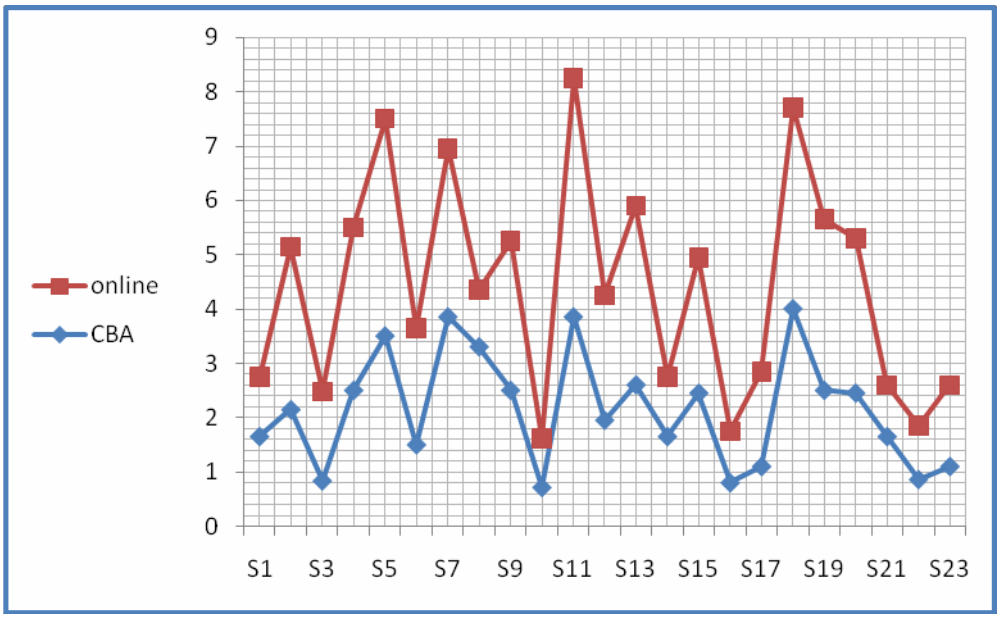


Figure 4-14 Subjective Usability of E-Assessments for CE416

Note that in Figure (4-14): CE416 online assessment was better than CBA offline, from the students' standpoint.

Table 4-9(a): Results for Subjective Metrics CN281

Subjective Code	Subjective Metrics	Question No	CBAs Mean	Online Mean
S1	Rating scale for multimedia usage	27,28	80% 4	60% 3
S2	Rating scale for appropriate feedback	8,5	25% 1.25	40% 2
S3	Rating scale for navigation	11	40% 2	60% 3
S4	Rating scale for the main menu	10	67% 3.35	73% 3.65
S5	Rating scale for pedagogic feedback	6	13.3% 0.665	33% 1.65
S6	Rating scale for easy to understand output	1	26.7% 1.335	40% 2
S7	Rating scale for loading application	4	46.7% 2.335	33.3% 1.665
S8	Rating scale for time to respond	3	80% 4	60% 3
S9	Rating scale for task effort	17,18	13.3% 0.665	20% 1
S10	Rating scale for finding help	13	36% 1.8	10% 0.5
S11	Rating scale for appropriate language	15	35.03% 1.7515	11.06% 0.553
S12	Rating scale for appropriate content	16	22.7% 1.135	29.5% 1.475
S13	Rating scale for ease of learning	19	71% 3.55	75% 3.75
S14	Rating scale for suitability for all users	21	64% 3.2	82% 4.1

Table 4-9(b): Results for Subjective Metrics CN281

Subjective Code	Subjective Metrics	Question No	CBAs Mean	Online Mean
S15	Rating scale for performance assessment	23	61% 3.05	84% 4.2
S16	Rating scale for error messages	24	71% 3.55	93% 4.65
S17	Ease of readability	25	84% 4.2	78% 3.9
S18	Satisfaction with text	26	76% 3.8	67% 3.35
S19	Rating scale for engagement	34,35	64% 3.2	76 % 3.8
S20	Rating scale for screen layout	29,30	57% 2.85	71% 3.55
S21	Rating scale for attractive screen design	31	73% 3.65	77% 3.85
S22	Rating scale for interface color	32	75% 3.75	79% 3.95
S23	Rating scale for icons and buttons	33	69.6% 3.48	79.8% 2.49

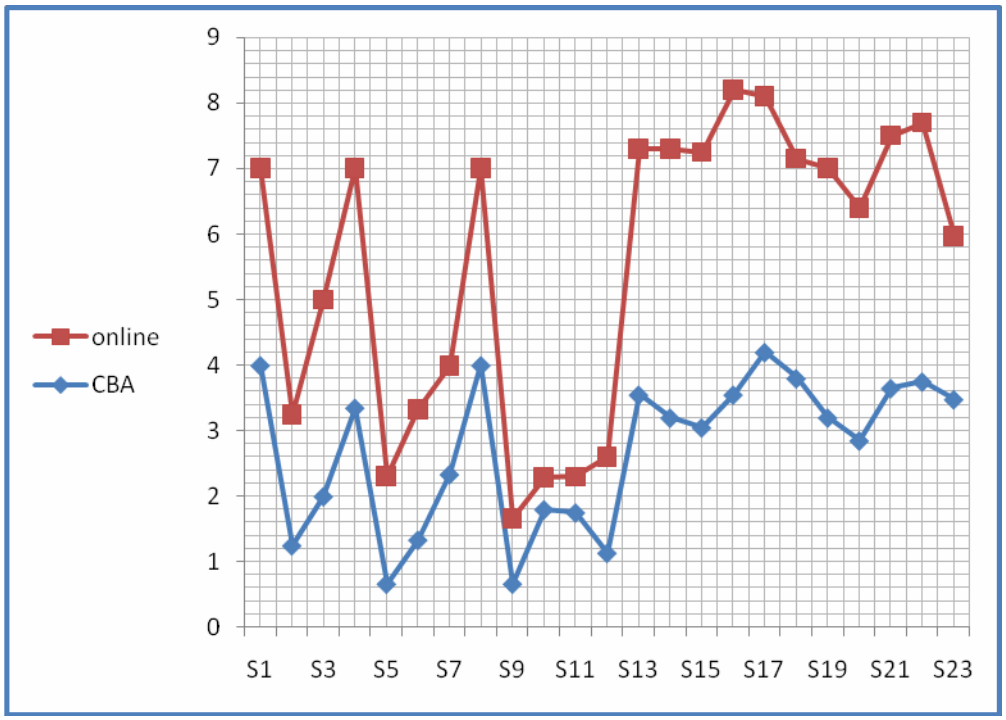


Figure 4-15 Subjective Usability of E- Assessments for CN281

Noted that in figure (4-15): In CN281 online assessment was better than CBAs offline, results were too convergent from students standpoint.

Table 4-10(a): Results for Subjective Metrics IS361

Subjective Code	Subjective Metrics	Question No	CBAs Mean	Online Mean
S1	Rating scale for multimedia usage	27,28	18% 0.9	12% 0.6
S2	Rating scale for appropriate feedback	8,5	41% 2.05	57% 2.85
S3	Rating scale for navigation	11	12% 0.6	20% 1
S4	Rating scale for the main menu	10	61% 3.05	69% 3.45
S5	Rating scale for pedagogic feedback	6	25% 1.25	34.1% 1.705
S6	Rating scale for easy to understand output	1	33.6% 1.68	47% 2.35
S7	Rating scale for loading application	4	87% 4.35	92% 4.5
S8	Rating scale for time to respond	3	32% 1.6	46% 2.3
S9	Rating scale for task effort	17,18	74% 3.7	89% 4.45
S10	Rating scale for finding help	13	17% 0.85	12% 0.6
S11	Rating scale for appropriate language	15	10% 0.5	23% 1.15
S12	Rating scale for appropriate content	16	80% 4	85% 4.25
S13	Rating scale for ease of learning	19	70% 3.5	70% 3.5

Table 4-10(b): Results for Subjective Metrics IS361

Subjective Code	Subjective Metrics	Question No	CBAs Mean	Online Mean
S14	Rating scale for suitability for all users	21	94% 4.7	79% 3.95
S15	Rating scale for performance assessment	23	11% 0.55	21% 1.05
S16	Rating scale for error messages	24	33% 1.65	44% 2.2
S17	Ease of readability	25	80% 4	75% 3.75
S18	Satisfaction with text	26	60% 3	50% 2.5
S19	Rating scale for engagement	34,35	80% 4	70% 3.5
S20	Rating scale for screen layout	29,30	95% 4.75	89% 4.45
S21	Rating scale for attractive screen design	31	88% 4.4	93% 4.65
S22	Rating scale for interface color	32	98% 4.9	79% 3.95
S23	Rating scale for icons and buttons	33	95% 4.75	69% 3.45

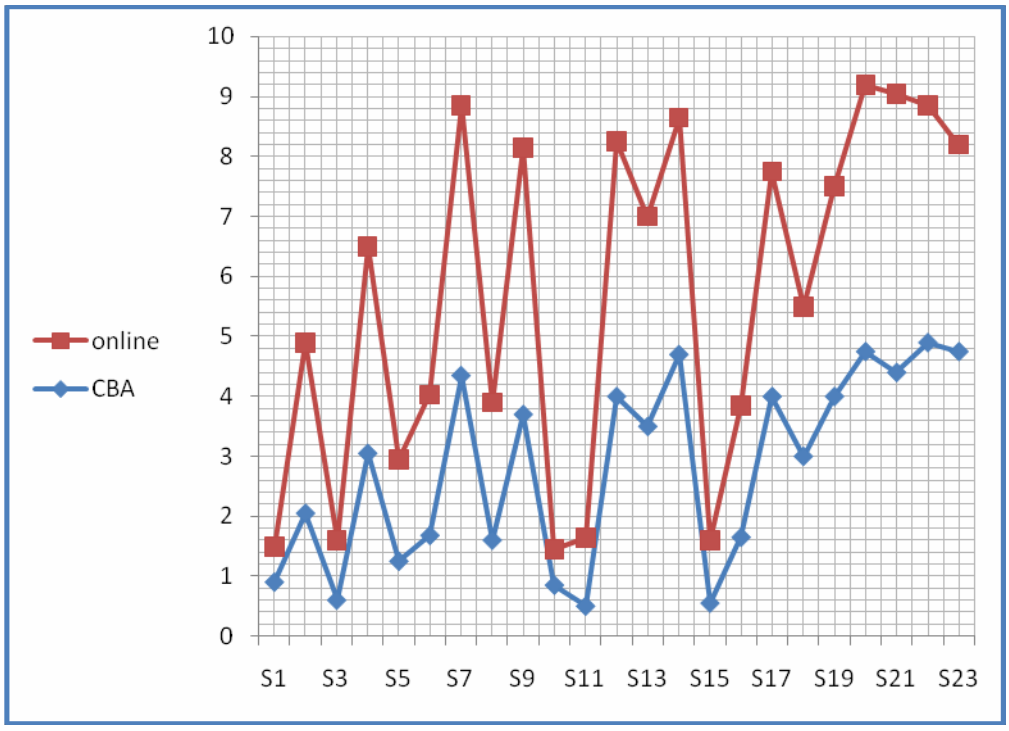


Figure 4-16 Subjective Usability of E- Assessments for IS361

Note that in Figure (4-16) : IS361 CBAs offline was better than online assessment, results were convergent from students standpoint.

Table 4-11(a): Results for Subjective Metrics IT112

Subjective Code	Subjective Metrics	Question No	CBAs Mean	Online Mean
S1	Rating scale for multimedia usage	27,28	27,28	10% 0.5
S2	Rating scale for appropriate feedback	8,5	8,5	23% 1.15
S3	Rating scale for navigation	11	11	64% 3.2
S4	Rating scale for the main menu	10	10	50% 2.5
S5	Rating scale for pedagogic feedback	6	25% 1.25	34.1% 1.705
S6	Rating scale for easy to understand output	1	6	66% 3.3
S7	Rating scale for loading application	4	1	84% 4.2
S8	Rating scale for time to respond	3	4	66% 3.3
S9	Rating scale for task effort	17,18	3	47% 2.35
S10	Rating scale for finding help	13	17,18	14% 0.7
S11	Rating scale for appropriate language	15	60% 3	40% 2
S12	Rating scale for appropriate content	16	70% 3.5	66% 3.3

Table 4-10(b): Results for Subjective Metrics IT112

Subjective Code	Subjective Metrics	Question No	CBAs Mean	Online Mean
S13	Rating scale for ease of learning	19	70% 3.5	75% 3.75
S14	Rating scale for suitability for all users	21	80% 4	90% 4.5
S15	Rating scale for performance assessment	23	77% 3.85	60% 3
S16	Rating scale for error messages	24	10% 0.5	26% 1.3
S17	Ease of readability	25	95% 4.75	70% 3.5
S18	Satisfaction with text	26	70% 3.5	75% 3.75
S19	Rating scale for engagement	34,35	50% 2.5	50% 2.5
S20	Rating scale for screen layout	29,30	95% 4.75	83% 4.15
S21	Rating scale for attractive screen design	31	85% 4.25	60% 3
S22	Rating scale for interface color	32	90% 4.5	88% 4.4
S23	Rating scale for icons and buttons	33	91% 4.55	80% 4

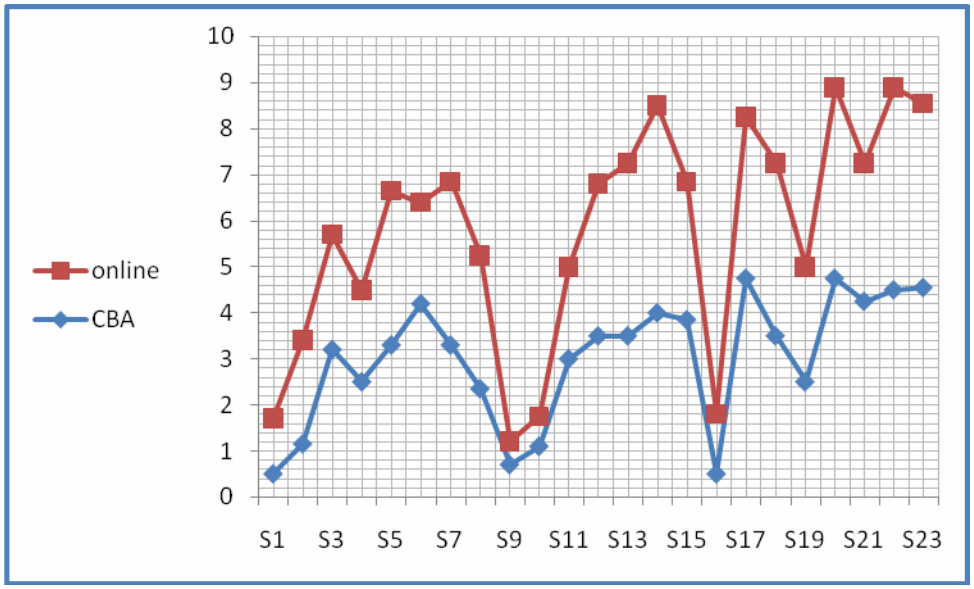


Figure 4-16 Subjective Usability of E- Assessments for IT112

Note that in Figure (4-15): IT112 CBAs offline was better than the online assessment; the results were too convergent from the student's standpoint.

4.8 Results Analysis

Comparative analysis introduces the results for subjective measures are presented diagrammatically from Figure(4-13) until Figure(4-16) for mean and percent respectively. The results of the subjective Usability for CBAs metrics express better subjective Usability than online assessments, although, the results too convergent, except for text size, the manner of questions view, performance assessment from the tutors' standpoint where online assessments showed better results, which means that students were more satisfied with online assessments and had good experience using it. However, both E- Assessments showed poor Usability help, and error messages. Furthermore, the students were unsatisfied with text size, readability of the CBAs. These user interface design attributes need to be improved.

The results indicate that the user satisfaction questionnaire developed in the model is reliable and effective for collecting subjective data for evaluating the Usability of E-Assessments.

The overall analysis shows that both subjective and objective results correlate, and closely linked harmonization between them. In addition to, the results showed that the model is not only useful for evaluating Usability and comparison of different E-Assessments but also helpful to uncover Usability issues and UX. To conclude, it is evident from the results that the MMB-UM model proposed in this research is effective and reliable.

4.9 Conclusion

The methodology which implemented of these experiments were presented in the previous chapter.In this chapter, the objectives, tools, hypotheses, and definition of the variables were determined. The experiments were conducted and the results were calculated and analysed.

Chapter5

Conclusion and future research

5.1 Research contributions

The contributions, which presented in this dissertation, as follows:

1. Connective between Human-Computer Interaction (HCI) and Electronic Assessments (E-Assessments) and combine them in one application of (HCI), which is the proposed model of Computer-Based Assessments (CBAs).
2. Presenting the MMB-UM model to measure the E-Assessments.

5.2 Discussion

In this research, the MMB-UM model has been used to measure E-Assessments, then validation of this model by applying the MMB-UM to a measure the Usability and the UX. The main purpose to measure the Usability and the UX by determining whether the model is effective to collect subjective(UX) and objective(Usability) data for both of them, analyses and compare the E-Assessments, to present the results, and highlighting the areas of improvement. The measurement can be modified by developing a new target or adding new questions. Thus, the developmental tasks and questionnaire can also be updated accordingly.

Through the experiments of CBAs offline, 17 students they were active and enthusiastic to do the experiments, and their observations were on the design and timing of each question, also, they had a problem in determining the time to answer each question and the problem of displaying one question at once,while the online assessments, the questions were presented in one go, nonetheless, the online assessments interface was not interesting, and forced them to answer all the questions.

The view which obtained from some the tutors were they satisfied with the E-Assessments when the experiments were conducted to compare between the E-

Assessments the Paper-Based Assessments(PBAs), the difference appeared at the time, the cost, and the effort were observed. Whereas some tutors against of E-Assessments for the following reasons :

- A number of security-related problems.
- The same exam might be used one more time.

In addition, they could not disregard the benefits of E-Assessments in:

- The risk of cheating is reduced or eliminated because every student has a different copy of the exam.
- The time for the tutors is saved by the formative assessments online, where each student can answer the exams from home.

Overall, the E-Assessments are successful in the students' who participated in the experiments because they found interest and pleasure.

5.3 Conclusions

The CBAs has become common in the Middle East Universities since 2010; also, Libya trended toward the electronic-education initiatives such as tests of computer driving license online and Faculty of Medicine at Benghazi University in 2012 by applying the correction based on the computer,. Etc.

The CBAs one of the HCI applications, which attract users its.The Usability and the UX are important factors to measure the HCI applications. Concepts of evaluation, Usability, and Usability testing are different and nested names, but provided the same results to measure Usability. Through this research, found that; the work of the E-Assessments requires to the existence of a reliable work team to enter data, safety is the most important condition with suitable the environment to execute.

In Libya, still suffer from the lack of appropriate environment of E- Assessments like lab equipment... etc., as well as the Internet.In the absence of all these requirements, the experiments were done by simple efforts and tools; also, faced many challenges are endless for example, tutors refusal to cooperate by giving the exam questions, despite their admiration with this idea, they describe it as time-saving, effort, and the cost reverses the PBA, also neither need for the presence of supervisors for the exams.

According to Faneer in (2015) theorists have tried to eliminate the borders between UX and Usability, others have tried to differentiate between them Vermeerenet al (2010) But differences of opinion still exist and research is still ongoing.

Must be noted that through the research and the experiments were applied in this research that: Ease of use represents the functionality, and user satisfaction represents the requirements of the user of the functionality. For example, if you asked for a car to move from Benghazi to Tripoli, the usability would be the car that arrived in Tripoli, and user experience is the comfort, safety, seats, and shape of the car, etc.

Through the previous literature, concluded that: the Usability is the objective that measures the quality of the E-Assessments, the satisfaction is the factor of the Usability, whereas UX measures user satisfaction of the E-Assessments. Usability consists of the satisfaction factor, which is a small part of the UX.

5.4 Recommendations for Future research

In the future, might have a workshop for creating suit environment full E-Assessments based on the group of the analysts and the designers who are executing the examinations. The E-Assessments can be applied to the sample of students include all the faculties of the University of Benghazi, therefore, can be applied to make sure its effectiveness through used it for one semester. The University can conclude contracts with software companies to create CBAs and purchase hosting online from trusted companies. The MMB-UM measurement, which presented in this research could be update and developed to match the changing of technology.

Reference

Algahtani, A.(2015) *The effect of multimodal metaphors on usability and communication performance of e-assessment*. Doctoral thesis . University of De Montfort.

Alshamari, M., & Mayhew, P. (2009). Current Issues of Usability Testing. . *IEEE Technical Review, 26*, 402-406.

Benyon D., Turner P., & Turner S. (2005) . *Designing interactive systems: a comprehensive guide to HCI and interaction design*, (2nd Ed). Pearson Education Limited, Edinburgh

Boone, H .N., & Boone ,D A.(2012)."Analyzing Likert data." *Journal of Extension, 50*,(2).

Booth P (1989). *An introduction to human-computer interaction*. Lawrence Erlbaum Associates Publishers, Hove/East Sussex.

Brown, J.D.(1997) Computer in Language testing ; present research and some future directions. *Language Learning and Technology, 1*,(1), 44-59.

Carroll, JM. (2002) .*Human-computer interaction in the new millennium*. Addison-Wesley, New York

Derom, J.(2015,August 13).*What is User Experience?*. Retrieved from https: //www. User testing .com /blog /author /jw/.

Dix, A., Finlay, J., Abowd, G., & Beale, R. (2004). *Human-Computer Interaction*. (3rdEd.). Harlow, Essex: Pearson Education Limited

Donald, D. 1998. *The design of everyday things.* London: MIT Press.

Faneer,M. (2015).*MULTIMODALE-LEARNING :AN EMPIRICAL STUDY.* DeMONTfort University press.

Harrington,HJ.RetirevedJanuary16,2017.from http://www.goodreads.com/quotes/42617.H_James_Harrington

Holbrook, N. J. , Devonshire, E. (2005) .Simulating scientific thinking online: An example of research-led teaching. *Higher Education Research & Development. 24* , (3), 201-213.

Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies, 64*, (2), 79-102.

ISO 9126-1:2001. *Software engineering - Product quality - Part 1:Quality model*. ISO, Geneva, Switzerland.

ISO 9241 -11 .(1998) .*Ergonomic requirements for office work with visual display terminals (VDTs) Part 11:Guidance on Usability*. ISO.

ISO 9241-210:2010. *Ergonomics of human-system interaction - Part210: Human-centred design for interactive systems*. ISO, Geneva,Switzerland.

Kabir,Md.,Han, b.(2016 August9). An Improved Usability Evaluation Model for point –of-Sale systemys. *International Journal of Smart Home, 10,* 269282 .doi: 10.14257/ijsh.

Kothari, C.R. (1990) .*Research Methodology: Methods and Techniques*. (2nd Ed.). New Age International Publisher.

Latour,T.,Sarre,S.(2018). ICASE Project:New Challenges In Computer –Based Assessment. ERCIM News Transfer.Retervied from http://www.tago lu.

Lynch,M.(2002*).The Online Educator: A Guide to Creating the Virtual Classroom*. Routledge

Matraf, M., Hussain ,A.(2017) .*Usability Evaluation Model for mobile e-book application*. Paper presented at Conference Proceedings**1891**, 020055 https://doi.org/ 10.1063/1.5005388.

McCall, J., Richards, P, Walters, G. (1977). *Factors in Software Quality, Nat'l Tech. Information Service*, no.Vol.1, 2 and 3.

Morgan, K ., Spector ,J.M .(2004).*The internet Society: Advances in Learning Commerce and Society* . Wit Press :Boston .

Nielsen, J., & Mack, R. L. (Eds). (1994). *Usability inspection methods*. NY.: John Wiley & Sons, Inc.

Nielsen, J. (1993). *Usability Engineering*. Academic Press, Boston, ISBN 0-12-518405-0.

Nielsen, J. (2003). Usability 101. Retrieved .December12,2017 from http://www.useit.com/alertbox/20030825.html .

Nottingham Trent University 2016 Retrieved from https://now.ntu.ac.uk

Palomba, C., W.Banta, T. 1999. *Assessment Essentials Planning Implementing and Improving Assessment in Higher Education* .Ed3-5. An Francisco: Jossey bass.

Palomba,C., & Banta, TW.,(1999) *Assessment Essentials Planning Implementing and Improving Assessment in Higher Education* .ed3-An Francisco: Jossey bass.

Plantak Vukovac, D., Kirinic,V.,& Klicek, B. (2010). *A Comparison of Usability Evaluation Methods for e-Learning Systems*, Chapter 27 in DAAAM International Scientific Book 2010, 271 -288, B. Katalinic (Ed.), Published by DAAAM International, ISBN 978-3-901509-74-2, ISSN 1726-9687, Vienna, Austria .DOI: 10.2507/daaam.scibook.2010.27

Preece, J., Rogers, Y., & Sharp, H. (2002). *Interaction design: Beyond human-computer interaction*. New York, NY: J. Wiley & Sons.

Preece, J., Rogers, Y., Sharp, H., Benyon,D., Holland, S., & CAREY, **T.** (1994) . *Human-Computer Interaction*. Wokingham, UK: Addison-Wesley.

Preece, J., Y. Rogers, et al. (1994). *Human-Computer Interaction*, Addison-Wesley.

Proprofs Training Maker.(2016,January16). What is Computer Based Assessments. Retrieved February 17,2018 from https://www. Proprofs .com lc/gay-5771182-digital/Independent and dependent Variable Examples.(nd).Retrieved December3,2017,from http://examples.your dictionary.com / independent-and-dependent- Variable –example .html.

Rajanen , D. , Clemmensen , I. , Livari, N, Inal ,Y.,Kizvaraglu, K. Sivayi, A., & Roche , A.(2017).*UX Definition of Usability XI – A Comparison Between Turkey, Finland , Denmark, France and Malaysia* . Paper presented at In IF IP Conference Human – Computer Interaction.216-239.Springer, Cham.

Seffah, A., Donyaee, M., Kline, R, B., &Padda H,K. (2010)."*Usability measurement and metrics"*.

Shackel,B. (1991). *Usability - context, framework, design and evaluation*. Shackel,b ., Richardson,S (Eds.), *Human factors for informatics usability*. 21-38. Cambridge: Cambridge University Press.

Shawgi ,E., Noureldien,A. Usability Measurement Model (UMM):Anew Model Measuring Websites Usability, *International Journal of Information Science ,5*,(1),5-13 .doi:10.5923/j.ijis.2015 0501.02.

Shneiderman, B .(1992). *Designing the User Interface: Strategies for Effective Human—Computer Interaction*,(2nd), Addison—Wesley.

Sim, G (2009) *Evidence Based Design of Heuristics: Usability and computer* Assisted Assessment .Doctoral thesis, University of Central Lancashire.

Simin,S.,Heidari,A. (2013,June 29).Computer Based assessment :pro and Cons. *Elixir International Journal,55,*12732-12734.Retrieved from http://www.elixerpublisher

Smith, M. S., Broom, M. (2003). *The landscape and future of the use of technology in K-12 education*. H. F. O'Neil, Jr., Perez, R. S. (Eds.), *Technology applications in education: A learning view*. 3-3. Mahwah, NJ: Erlbaum

Spool, J.(2007) *The difference between usability and user experience*. Retrieved from http:www.uie.com/brain Sparks/2007/03116/the-difference –between usability –and-user-experience, accessed 4th August 2015.

Tahir,R. &Arif, F. (2014) .A measurement Model Based on Usability Metrics of Mobile Learing user interface. The society of information and wireless communication.

Tahir,R. &Arif,F.(2014) A measurement Model Based on Usability Metrics of Mobile Learing user inter face . The society of information and wirlesse communication.

Thelweall, M.(2000). Computer –Based Assessment : a versatile educational tool .*EDerierscience Gd*, 34,37-49.

Valenti,s., Cucchiarelli,A.,&Panti,M.(2003). Computer Based Assessment Systems Evaluation Via the ISO 9126 Quality Model. INK.Nantz (Ed) .*Journal of Information Technology Education,1*,(3).

Vermeeren., Arnold POS., Law., Virpi., Obrist., Hoonhout.,&Väänänen-Vainio-Mattila.(2010). "User experience evaluation methods: current state and development

needs."presented at the 6th Nordic Conference on Human Computer Interaction: Extending Boundaries,. 521-530. ACM.

Wiles,A .(2015) Library usability in higher education: how user experience can form library policy . Doctoral thesis, University of West London.

Wiles,Alison.(2015). *Library usability in higher education: how user experience can form library policy*. Doctoral thesis, University of West London.

www.King Abdul Aziz University.com Retrieved on 5/9/2016

www.usability.com Retrieved on 6/9/2016

Yang, X .,Chen,G .(2009). *Human Computer Interaction Design in product Design* . University of Wuhan.

Bibliography

- Lew, P., Zhang, L. and Wang, S.*' Model of Measurement for Web Application: from an End User Perspective*, China: Beilnang University.
- Mackenzie, L., and Ballarnd, K. (2015) Can use Individual Online Interactive Activities Enhance Exam Result? *MERLOT Journal of Online Learning and Teaching*, *11*(2), 262-266.
- Muungi, J. and Tossy, T.' Usability Evaluation Methods and Principles for The Web, *International Journal of Computer Science of Soft Ware Engineering (IJCSSE)*, 4(7), 165-171.
- Harper, R., Rodden, T, Rogers, Y .and Sellen, A. (Eds.). (2008).*' Being Human: Human Computer Interaction in The Year 2020.*England: Microsoft research Ltd.
- Valenti, S., Cucchiarell, I., & Panti, M. (2002) . ' Computer Based Assessment Systems Evaluation Via the ISO 9126 Quality Model ,*Journal of Information Technology Education*, *1*(3).

APPENDICES

APPENDICE A

1. Usability Criteria

•Learn _ability: by which new users can begin effective interaction and achieve maximal performance;
•Flexibility: the multiplicity of ways the user and system exchange information;
•Robustness: the level of support provided to the user in determining successful achievement and assessment of goals;
•Efficiency: once the user learns about the system,[the speed with which s/he] can perform the tasks;
•Memory _ ability: how easily the user will remember the system functions, after a period time of not using it;
•Errors: “How many errors do users make, how severe are these errors, and how easily can they recover from the errors?” (Nielsen 2003);
•Satisfaction: how enjoyable and pleasant is it to work with the system?".

2. Usability Inquiry Approach

1. **Field Observation**
    - Go to user work places and observe them work
    - How the users are using the system to accomplish their tasks
    - what kind of mental model the users have about the system
2. **Focus Groups**
    - Data collecting technique where about 6 to 9users are brought together to discuss issues relating to the system.
    - A human factors engineer plays the role of a moderator, who needs to prepare the list of issues to be discussed.
3. **Interviews**
    - Human factors engineers formulate questions about the product based on the kind of issues of interest.
    - Then they interview representative users to ask them these questions in order to gather information desired.

- In an evaluation interview, an interviewer reads the questions to the user, the user replies verbally, and the interviewer records those responses.
- Interviews can be Structured & Unstructured

4. **Unstructured**
    - During early stages of usability evaluation.
    - Does not have well defined agenda, and is not concerned with specific aspect of the system.
    - Obtain information on procedures adopted by users and on their expectations of the system.

5. **Structured**
    - Has a specific, predetermined agenda with specific questions to guide and direct the interview.

6. Questionnaire

- This is the method, being used for a long time.

APPENDICE B

1. Computer Based Assessment Definition (CBA)

Many researchers and international miens try to define" assessment" word; definitions of assessment

Palomba and Banta (1999). :"Assessment is the systematic collection, review, and use of information about educational programs undertaken for the purpose of improving learning and development". The Higher Learning Commission provides a definition for assessment as: " Assessment of student learning is a participator, iterative process that: provides data/ information you need for your students learning. Engages you and others in analysing and using this data/information to confirm and improve teaching and learning produces evidence that students are learning the outcomes you intended guides you in making educational and institutional improvements evaluates whether changes made improve/impact student learning, and documents the learning and your efforts". University of Oregon, Teaching Effectiveness Program: " we define assessment as follows: Assessment is the process of gathering and discussing information from multiple and diverse sources in order to develop a deep understand ,

and can do with their knowledge as a result of their educational experiences , the process culminates when assessment results are used to improve subsequent learning" .

According to Simin and Heidari (2013)

2. Administrative Advantages/ Disadvantages of Computer Based Assessments(CBA)

- **Administrative advantages**

1. Computerized marking is not prone to human error.
2. Saves staff time in terms of supervising and marking (including double marking) assessments.
3. Reduction of printing costs, particularly when tests are updated or changed.

- **Administrative disadvantages**

1. Implementing a CBT can be costly and time-consuming especially when trying to integrate with an institution's MLE.
2. Staffs who design and invigilate CBT need training in assessment principles and design, IT skills and examination management.
3. A high level of collaboration between all those involved in designing and implementing CBT required.
4. Some systems cannot implement anonymous marking.
5. Hardware and software used to deliver CBT needs to be robust in order to avoid failure at crucial times such as examinations.

3. Pedagogical Advantages/ Disadvantages

- **Pedagogical advantages**

1. Tutors can incorporate hints into test questions.
2. Tutors can monitor the progress of students through frequent use of assessment.
3. Students can monitor their own progress and revise and rehearse at their own pace.

4. Detailed and specific feedback can be given to students during and immediately after a test.
5. Tutors can assign different learning activities to students based on their test results.
6. Can provide tutors with feedback for evaluation of modules/courses/programs.

- **Pedagogical disadvantages**

1. Unsupervised CBT sessions present a risk of plagiarism (it can be difficult to authenticate the identity of students).
2. Students need to have sufficient IT skills and experience of the requirement of CBT.
3. Staffs have a tendency to just use MCQs which can be tedious and motivating for students, and it has been argued that MCQ focus on testing superficial levels of students learning.

Other advantages

Timely feedback; the teacher can provide feedback.

1. Automatic feedback; some forms of on-line assessment answers (i.e. multiple choices).
2. Monitoring and tracking of learners' results behavior.
3. Choice of assessment modes, such as multimedia, interactivity, etc.
4. Time-saving; an assessment can be created using software tools and adapted and reused as needed. They can be distributed and collected using a web-based system which saves development and distribution.
5. Reduces resources needed by replacing human resources with computer resources.
6. Reduces turnaround time; as the systems enables assessments to be corrected by computers. Reduced time further enables students to use the knowledge obtained from corrected assessment to address further assessment sooner.
7. Keeping records of results that can be stored centrally and assessed by interested parties, such as students and staff.

8. Increasing ease with which data can be used as corrected assignment corrected and stored electronically can be analyzed easier and the data can be used in spreadsheets and other statistical packages.
9. Flexible and comfortable environment; on-line tests afford students the opportunity to take tests on their own terms.
10. Time-consuming grading can be done by assessment software.
11. Once taken and graded can be reconfigured for multiple attempts, providing practice tool for students.
12. Computers are more accurate at scoring selected-response tests than human beings are.
13. Computers are more accurate at reporting.
14. Computers can give immediate feedback.
15. Diagnostic feedback can be provided very quickly to each students on those items answered incorrectly if that is the purpose of the test.

- **Other disadvantages**

1. A high level of organization is required across all parties involved in assessment (academics, support, staff, computer services and administrators).
2. Assessors and invigilators need training in assessment design, IT skills and examination management.
3. Hardware and software must be carefully monitored to avoid failure during examination, and students require adequate IT skills and experiences of the assessment type.
4. Construction of good objective tests requires skills and practice and so is initially time consuming and because of this, testing of higher other skills is difficult.. Computer anxiety.
5. Differences in the degree to which students are familiar with using computers.
6. Technical malfunctions; computer equipment may not always be available or in working order.
7. Cheating will arise.

8. Absence of instructor; instructor is not on-demand when has a question about a problem or when a student may be confused by the language of the problem.

APPENDICE C

The ISO/IEC 9126-4 approach to Usability Metrics

The ISO 9241-11 standard defines usability as "the extent to which a product can be used by specified users to achieve specified goals with **effectiveness**, **efficiency** and **satisfaction** in a specified context of use". The reason why I marked effectiveness, efficiency and satisfaction in bold is because this definition clearly states that usability is not a single, one-dimensional property but rather a combination of factors.

The ISO/IEC 9126-4 Metrics recommends that usability metrics should include:

- **Effectiveness:** The accuracy and completeness with which users achieve specified goals
- **Efficiency:** The resources expended in relation to the accuracy and completeness with which users achieve goals.
- **Satisfaction:** The comfort and acceptability of use.

However, the actual ways of how these should be measured is very often left at the discretion of the evaluator.

1. Usability Metrics for Effectiveness

1.1 – Completion Rate

Effectiveness can be calculated by measuring the **completion rate**. Referred to as the fundamental usability metric, the completion rate is calculated by assigning a binary value of „1" if the test participant manages to complete a task and „0" if he/she does not.

Due to its simplicity, the completion rate is a metric that is very easy to understand, hence the reason why it is very popular. Moreover, it can be collected during any stage of development. Effectiveness can thus be represented as a percentage by using this simple equation:

$$Effectiveness = \frac{\text{Number of tasks completed successfully}}{\text{Total number of tasks undertaken}} \times 100\%$$

1.2 – Number of Errors

Another measurement involves counting the number of errors the participant makes when attempting to complete a task. Errors can be unintended actions, slips, mistakes or omissions that a user makes while attempting a task. You should ideally assign a short description, a severity rating and classify each error under the respective category. Although it can be time consuming, counting the number of errors does provide excellent diagnostic information.

Based on an analysis of 719 tasks performed using consumer and business software, Jeff Sauro concluded that the average number of **errors per task is 0.7**, with 2 out of every 3 users making an error. Only 10% of the observed tasks were performed without any errors, thus leading to the conclusion that it is perfectly normal for users to make errors when performing tasks.

2. Usability Metrics for Efficiency

Efficiency is **measured in terms of task time**. that is, the time (in seconds and/or minutes) the participant takes to successfully complete a task. The time taken to complete a task can then be calculated by simply subtracting the start time from the end time as shown in the equation below:

Task Time = End Time – Start Time

Efficiency can then be calculated in one of 2 ways:

2.1 – Time-Based Efficiency

$$Time\ Based\ Efficiency = \frac{\sum_{j=1}^{R}\sum_{i=1}^{N}\frac{n_{ij}}{t_{ij}}}{NR}$$

Where:

N = The total number of tasks (goals)

R = The number of users

n_{ij} = The result of task i by user j; if the user successfully completes the task, then

N_{ij} = 1, if not, then N_{ij} = 0

t_{ij} = The time spent by user j to complete task i. If the task is not successfully .completed, then time is measured till the moment the user quits the task.

2.2 – Overall Relative Efficiency

The overall relative efficiency uses the ratio of the time taken by the users who successfully completed the task in relation to the total time taken by all users. The equation can thus be represented as follows:

$$Overall\ Relative\ Efficiency = \frac{\sum_{j=1}^{R}\sum_{i=1}^{N} n_{ij}t_{ij}}{\sum_{j=1}^{R}\sum_{i=1}^{N} t_{ij}} \times 100\%$$

3. Usability Metrics for Satisfaction

User satisfaction is measured through **standardized satisfaction questionnaires** which can be administered after each task and/or after the usability testing session.

3.1 – Task Level Satisfaction

After users attempt a task (irrespective of whether they manage to achieve its goal or not), they should immediately be given a questionnaire so as to measure how difficult that task was. Typically consisting of up to 5 questions, these post-task questionnaires often take the form of Likert scale ratings and their goal is to provide insight into task difficulty as seen from the participants‟ perspective.

The most popular post-task questionnaires are:

- ASQ: After Scenario Questionnaire *(3 questions)*
- NASA-TLX: NASA‟s task load index is a measure of mental effort *(5 questions)*
- SMEQ: Subjective Mental Effort Questionnaire *(1 question)*
- UME: Usability Magnitude Estimation *(1 question)*
- SEQ: Single Ease Question *(1 question)*

From the above list, Sauro recommends using the SEQ since it is short and easy to respond to, administer, and score.

The Single Ease Question

3.2 – Test Level Satisfaction

Test Level Satisfaction is measured by giving a formalized questionnaire to each test participant **at the end of the test session**. This serves to measure their impression of the overall ease of use of the system being tested. For this purpose, the following questionnaires can be used (ranked in ascending order by number of questions):

- SUS: System Usability Scale *(10 questions)*
- SUPR-Q: Standardized User Experience Percentile Rank Questionnaire *(13 questions)*
- CSUQ: Computer System Usability Questionnaire *(19 questions)*
- QUIS: Questionnaire For User Interaction Satisfaction *(24 questions)*
- SUMI: Software Usability Measurement Inventory *(50 questions)*

  The natural question that comes to mind is … "which questionnaire should I use?"

  Garcia states the choice depends on the:

- The budget allocated to measuring user satisfaction
- Importance that the user‟s perceived satisfaction has on the overall project

  In fact, he recommends that SUMI should be used if there is enough budget allocated and the users‟ satisfaction is very important. If the measurement of user satisfaction is important, but there is not a large allocated budget, then one should use SUS.

On the other hand, Sauro recommends using SUS to measure the user satisfaction with software, hardware and mobile devices while the SUPR-Q should be used for measuring test level satisfaction of websites. SUS is also favoured because has been found to give very accurate results. Moreover, it consists of a very easy scale that is simple to administer to participants, thus making it ideal for usage with small sample sizes.

Usability& UX

1 I found it easy to understand

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

2 The application is too slow i had to wait for response to continue

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

3 The app took a lot of time for loading

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

4 The app provides a visual display to show the loading process

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

5 The app gives the feed back on whether my answer is correct or wrong

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

6 The application does not provide appropriate feed back for my answers

A	Poor	B	Fair		Neutral
D	Good	E	Excellent		

7 I was comfortable with the screen orientation of application

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

8 The main menu of application is confusing

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

9 The app provide clear and understandable navigation keys such as Back/next buttons to move to previous/next screen.

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

10 The application provides useful help information.

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

11 It was difficult to find help

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

12 It was difficult to understand the language used in the application

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

13 The topics /concept and information was understandable

A	Poor	B	Fair		Neutral
D	Good	E	Excellent		

14 I need to remember a lot of information throughout several actions to perform a task

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

15 It is was easy to complete the tasks without much effort

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

16 It is difficult to learn to use the application

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

17 The application provides different a progress report/ result for my performance in every activity

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

18 The application gives error messages that clearly tell me how to fix problems

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

19 It was easy to read the text in this application

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

20 The text size used in this application is too small

A	Poor	B	Fair		Neutral
D	Good	E	Excellent		

21 It is easy to find the information I needed

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

22 I find the design of application are not attractive

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

23 The colors used in this application are not attractive

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

24 The icons and buttons used are attractive and recognizable

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

25 The application gives interesting rewards on my performance

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

26 Overall i enjoyed using the app

A	Poor	B	Fair	C	Neutral
D	Good	E	Excellent		

الملخص

تؤثر الثورة العلمية الحديثة على النظام التعليمي، وخاصة في الجزء الحاسم من العملية التعليمية (التقييم)، يتكون التقييم التقليدي من بعض القضايا المعقدة المتعلقة بالوقت، والجهد، بالإضافة إلى تكلفة استهلاك الأوراق. تعتبر التقييمات الإلكترونية (التقييمات الإلكترونية) أحد تطبيقات التفاعل بين الإنسان والحاسوب لذلك، نستطيع قياس سهولة الاستخدام وتجربة المستخدم للتقييمات الالكترونية. في هذه الأطروحة، طبقنا نموذج مقترح للتقييمات الإلكترونية لأول مرة في كلية تقنية المعلومات في جامعة بنغازي. أجرينا تجربة التقييم الإلكتروني على مجموعة من الطلاب وبعض المعلمين في التقييم على الانترنت والتقييم باستخدام الحاسوب ثم، مقارنة النتائج بينهما وقياس العاملين: قابلية الاستخدام وتجربة المستخدم من خلال تصميم وتطبيق مقياس نموذج المستند على مقاييس قابلية الاستخدام لأنظمة الالكترونية.

نموذج القياس المعتمد على مقاييس سهولة الاستخدام لأنظمة الالكترونية (دراسة الحالة، كلية تقنية المعلومات)

قدمت من قبل:

إيمان عيسى صالح المسماري

إشراف:

د. كنز امحمد بوزيد الحسوني

قدمت هذه الرسالة استكمالا لمتطلبات الحصول على درجة الماجستير في هندسة البرمجيات.

جامعة بنغازي

كلية تقنية المعلومات

يوليو 2018