International Conference on Green Energy & Environmental Engineering (GEEE-2018)
Proceedings of Engineering and Technology – PET
Vol.38 pp.99-105

Usability Evaluation Model for Electronic Assessments

Eman I.B. Elmasmare[1], Kenz A. Bozed[2], Youssif H. Mousa[3]

Faculty of Information Technology, Benghazi University
Benghazi-Libya
[1]Eman.Elmasmare@uob.edu.ly
[2]Kenz.Bozed@uob.edu.ly
[3]Youssif.Mousa@uob.edu.ly

***Abstract*— In the last decade of the twentieth century and the beginning of the twenty-first century, an information technology has witnessed tremendous progress, that was led to a modern scientific revolution penetrated all disciplines and fields until it reached to the education, especially at critical part of education process it is(The assessment), whereby 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 UX for E-assessments. This paper will presented proposed model to evaluate usability for E-Assessments, to ensure reliability of the model, it was validated by applying the proposed model on two E-Assessments (Computer-Based Assessments, Online assessments).Results calculate and proved by A Measurement Model Based on Usability Metrics (MMB-UM) for E-Assessments.**

***Keywords*— Usability Models; Electronic Assessments (E-Assessments); Computer Based Assessments (CBAs); Usability evaluation; Measurement Model Based on Usability Metrics (MMB-UM) .**

I. INTRODUCTION

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, asked when describes the contradiction of technology "But what good is technology if it is very complicated to use"[1].

Usability cited as a concept appeared as a result of HCI and it is considered one of the most important quality factors for the web application [2]. Also, 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"[3]. Recently, contributed and indicated there is no a circumscribed definition for Usability and UX [4]. 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. [5],[6]).

II. RELATED WORK

A. Usability concept overview

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. 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"[6]. 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" [7]. "Usability is a quality attribute that assesses how easy user interfaces are to use" [8]. 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 1 summarized some of these definitional sets. Concerning attributes of Usability apparent in the table row, for example, all sources in Table 1 describe "efficiency" as Usability attribute, although not all sources use this particular term.

ISSN 2356-5608

Table 1. Sets of definitions specific attributes of Usability

Nielsen (1993)	**Constantine & Lockwood (1999)**	**Constantinos & Dan (2007)**
Efficiency of use	Efficiency in use	Efficiency
Learn ability (Ease of learning)	Learn-ability	----------
Memorability	Remember ability	----------
Errors/ Safety	**Reliability**	**Effectiveness**

.

B. Overview Usability Models

Usability concept has been defined in multiple ways, also evolved over a period [9]. There are three major to identify efficiency, effectiveness, and satisfaction as key attributes [3]. Usability as software quality attribute decomposed into five factors understandability, learn-ability, and operability, attractiveness, and Usability compliance [5]. There are many Usability models, but no one of these models covers all aspects of Usability. Such as Usability, models include [10].

C. 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 [7]. The Computer-Based Assessment can contain elements of the multi-media such as video, images, and sounds, different from the limited evaluation over Paper Based Assessments [11]. (PBAs) 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) 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 [6]. 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, the purpose of assessment is making decisions or judgments about the students, by the system of collecting the information [8]. 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 [12].

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

There are many famous measurements, models to measure both factors [13].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 provided , 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 [14]. By based they in this paper 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 Fig. 1 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 Usability factors for E-Assessments. The details are listed in Table 2 and each factor corresponds to a description, clarified 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.

ISSN 2356-5608

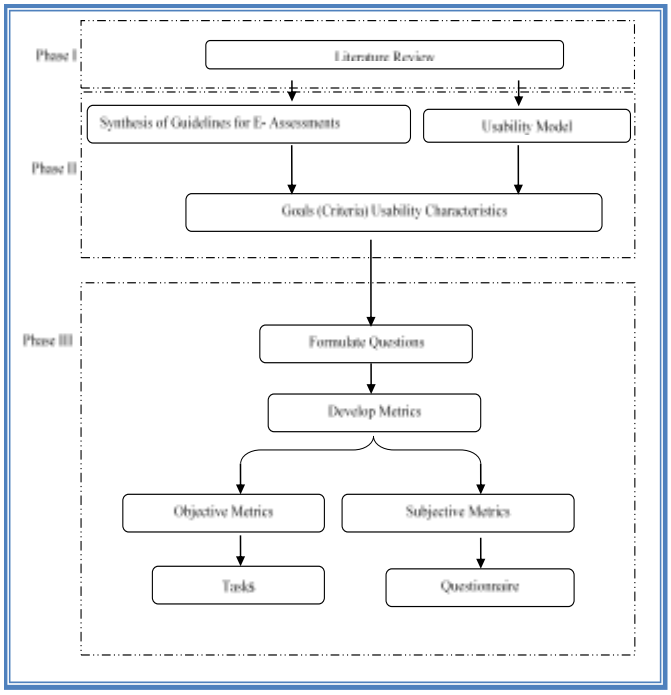

Fig. 1 General Structure of proposed Model (M M B- U M) for E-Assessments

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

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" [14]. As shown in Table 4 and Table 5. The questionnaire has 26 questions appropriate for E-Assessments, based on the Usability factors, which defined in Table 2.

Table 2. Proposed Usability Factors for E- Assessment software

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

Table 3. Sets of Phase II is illustrating Usability Characteristics, Goals, Questions and Metrics

#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.

ISSN 2356-5608

Table 4. Phase III Questionnaire of MMB_UM for E-Assessments

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
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
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 5. Phase III Task List of MMB-UM for E-Assessments

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
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**

IV. Experiment Results

Data compiled and analyzed according to answers to the questionnaire; it is given to twenty students of Information Technology-Faculty (ITF).

The experiments were conducted in a quiet and comfortable lab at Benghazi University. Before the

ISSN 2356-5608

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.

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 6) 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.

A. *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 7, which calculated by Usability metric equations, in [15] had indicated to (ISO/IEC 9126-4) that recommends that Usability metrics should include obtaining outcomes effectiveness, efficiency, and user satisfactions. 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 7 for both E- Assessments for -.

Table 6. Objective Usability Results of Experiments

Objective Code	Objective Metrics	CBAs Mean	On-line 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

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 Fig. 2, except for navigation.

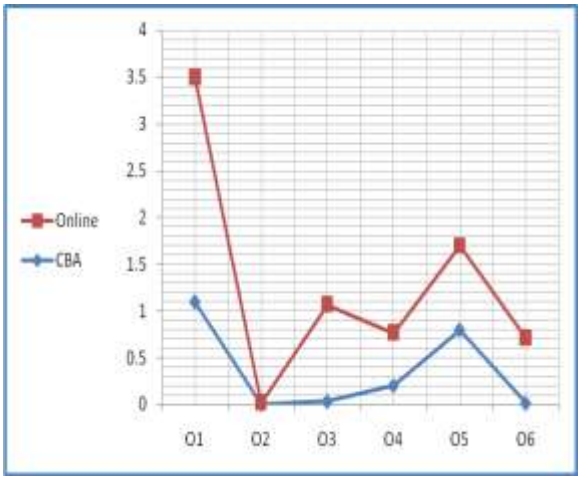

Fig. 2 General Objective Usability of E- Assessments for A-course

B. *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 satisfaction questionnaire was done according to the statistical procedure described [16]. For each subjective metric, the questions from satisfaction questionnaire were matched with the metric. In Table 7 presents the results of subjective measures for course-A. The labels "S1 to S23" are used to represent each subjective metrics.

Table 7(a). Subjective Usability Results of Experiments

Subjective Code	Subjective Metrics	Question No	CBA Mean	On-line 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

ISSN 2356-5608

Table 7(b). Subjective Usability Results of Experiments

Subjective Code	Subjective Metrics	Question No	CBA Mean	On-line 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
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
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

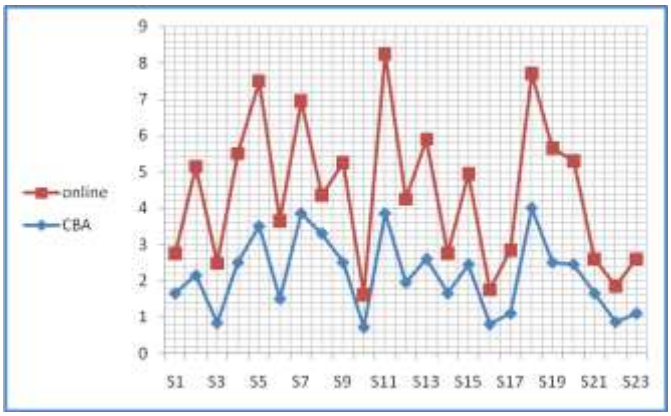

Fig. 3 General Subjective Usability of E- Assessments for A-course

Note that in Fig.3: A-course online assessment was better than CBA offline, from the students' standpoint.

ISSN 2356-5608

V. CONCLUSION

Comparative analysis introduces the results for subjective measures are presented diagrammatically from Figure 3 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.

The CBAs one of the HCI applications, which attract users its. The Usability and the UX are important factors to evaluate the HCI applications. Concepts of evaluation, Usability, and Usability testing are different and nested names, but provided the same results to measure Usability. 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. Finally this model can be develop and change.

REFERENCES

[1] D.Donald, The design of everyday things, London: MIT Prees, 1998.

[2] C. Palomba, and TW.Benta, 2nd Ed., Assessment in Higher Education, Anfrancisco: Jossey bass, 1999.

[3] Guidance on Usability, ISO Standard 9241-11 -1998(E).

[4] D. Rajanen, T. Clemmensen, N. Iivari, Y. Inal, K. Rızvanoğlu, A. Sivaji, and A. Roche, UX Professionals' Definitions of Usability and UX–A Comparison Between Turkey, Finland, Denmark, France and Malaysia.In IFIP Conference on Human-Computer Interaction, Springer, Cham, 2017, September, pp. 218-239.

[5] Software engineering- Product quality- Part 1: Quality model, ISO Standard 9126-1-2001.

[6] (2016) The Nottingham website. [Online]. Available: http://www.now.nut.ac.uk/

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

[8] M. Lynch, The Online Educator: A Guide to Creating the virtual Class Room. , Rout ledge, 2002.

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

[10] J. Nielson. (2003) Usability homepage on USEIT. [Online]. Available: http://www.useit.com/alertbox/20030825.html.

[11] K. Morgan and J.M. Spector, The internet Society: Advances in Learning Commerce and Society, Boston: WH Press, 2004.

[12] (2016) The Proprofs Training Maker website. [Online]. Available: 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.

[13] M. Matraf and A. Hussain, "Usability Evaluation Model for mobile e-book application", Paper presented at Conference Proceedings1891, 020055. Available: https://doi.org/ 10.1063/1.5005388, 2017.

[14] R. Tahir and F. Arif, "A Measurement Model Based on Usability Metrics of Mobile Learning User Interface", The society of information and wireless communication, 2014.

[15] (2016) The Usability website. [Online]. Available: http://www.Usability.com/

ISSN 2356-5608