Usability Evaluation Techniques: A Comprehensive Review

Dr. Leena Jain
Professor & Head-MCA, Global Institutes of Management & Emerging Technologies

Ramandeep Kaur
Research Scholar, IKG Punjab Technical University, Jalandhar, Punjab, India

Abstract: Usability is recognized as an important product attribute that affects the quality of software. Recently researchers have affirmed usability to be 'a necessary condition for survival'. Since usability is an amalgamation of a lot of fuzzy factors, it is very challenging to evaluate it in the 'real sense'. Though many usability evaluation techniques have been proposed so far but they are limited in scope and fail to cover the diverse and widespread dimensions of usability. This paper aims to compare and contrast a range of existing usability evaluation techniques with respect to their potentials, inadequacies and suitability for e- learning systems.

Keywords: evaluation; inspection; inquiry; testing; usability.

1. Introduction: The wave of usability largely started in 1980's, however; it gained significant importance by the year 1990. Several different usability evaluation methods (UEMs) were proposed, including Heuristic Evaluation, Cognitive Walkthroughs, Think-a-loud Protocol, Pluralistic Walkthrough, Surveys, Formal Inspections, etc. A lot of study in this field is led by a constant argument of proving the effectiveness of one method over the other. (Jeffries et al., 1991), (Karat et al., 1992), (Desurvire & Thomas, 1993),(Virzi et al., 1993)).Due to the lack of a single standardized comparison criteria (Lund, 1998), the researchers have compared the performances of these techniques through different measures. For example to test the usability of online library catalogues, Haak et al. (2004) compared two UEMs on the basis of usability problems detected; relevance of the problems identified; overall task performance; and participant experiences.Doubleday et al. (1997) assessed the UEMs on the criteria of problem focus; quality of the results and cost effectiveness. The comparative studies of UEMs differ not only in terms of their measuring criteria but also with respect to their goals. For example Molich et al. (2004) pointed out the results of a comparative evaluation of a web site-Microsoft Hotmail by nine independent teams of experts. The objective of their study was to examine the consistency of the results obtained. Hertzum and Jacobsen (2001) explored the 'evaluator effect' in usability evaluation with the intent of highlighting the fact that different evaluators evaluating the same interface with the same UEM identify different sets of problems. The general lack of understanding of the potentials and inadequacies of each of these methods has augmented the need to determine the objective, suitability and effectiveness of these UEMs.

2. Usability Evaluation Methods: There are many different ways in which usability evaluation methods can be classified. The following section elaborates these methods along with their strengths and weaknesses.

2.1 Inspection Methods:

2.1.1 Heuristic Evaluation: It is an informal usability evaluation technique where experts carry out a run through to judge the system design for its compliance to the established usability principles, called 'heuristics'. Heuristics are the general guidelines or rule of thumb that is used to steer a design decision or review an already taken decision. It is usually conducted by a group of experts as it is believed that a single

individual will not be able to identify all usability issues. Moreover, different experts tend to figure out usability problems from different perspectives leading to detection of wider set of problems. There are

numerous sets of heuristics documented in the literature namely Nielson's ten usability heuristics, Gerhardt-Powals 10 Cognitive Engineering Principles(Gerhardt-Powals,1996), Tognazzini's First Principles of Interaction Design (Tognazzini, 2003), Schneiderman's Eight Golden Rules of Interface Design(Schneiderman,1998), Arnold Lund's "Expert Ratings of Usability Maxims"(Lund, 1997). Practitioners over the years have adapted, extended and modified these predefined heuristics to better address the usability issues as per their diverse domain requirements. For example to design a child e-learning application, there was a need of more comprehensive heuristics pertaining to the requirements of a child (screen layout, challenge & curiosity) and of e-learning application (content design, assessment, interactivity)(Alsumait, A., and Al-Osaimi,2009). Similarly for designing a gaming application, aspects such as game play, game story, mechanics and usability needed to be included (Desurvire et al., 2004). For a website, design issues like Information Design, Consistency, Navigation, Operation and Errors were required to be addressed (Rau P, 2003). Thus literature is full of such examples of adapted heuristic sets developed to cater to domain specific usability needs.

Advantages	Disadvantages
It is most widely used usability evaluation technique in industry as well as academia(Collazos et al., 2017)	Results obtained are not reliable as it largely depends on the competence of the evaluator(Shoug,2015)
Its flexibility allows appraisals to be performed during early designs, prototypes, storyboards or even at the completion stage.	It tends to figure out many false positive usability problems i.e the problems that the actual user never faces(Zaharias and Koutsabasis, 2011)
Such evaluations are considered cheap as they do not involve the real users and the usability issues can be diagnosed early in the life cycle, much before the resource commitments.	Challenging to sum up the findings of multiple evaluators as each evaluator reports problems from his perspective.

2.1.2 Cognitive Walkthrough: Cognitive Walkthrough (CW) is the usability inspection method that stresses upon ease of learning. This evaluation technique is based on theory of learning by exploration and involves simulating the user behavior and actions when performing a specific task step by step and answering a set of questions at each step (Holzinger, 2005). The objective of this usability inspection method is to determine the usability issues and understanding the learnability of the system for new and novice learners. Cognitive walkthroughs are very cost effective and quick to perform as compared to other types of usability testing. They can be incorporated during the design phase i.e much prior to the development process and thus ensures resources are not spent on unusable solutions. The CW technique has undergone many variations to accommodate the need and time plans of the project. For example Rowley& Rhoades(1992) used video recording setups and informal interactions to study the user interfaces and termed the technique cognitive jogthroughs. Some of the advantages and disadvantages(Rowley& Rhoades, 1992),(Green & Burnett, 2000) of CW are

Advantages	Disadvantages
Can work even without functional model of the product.	Lacks guidelines about what makes an action clearly available to a user.
Cost effective and quick	Lack of information about what type of actions are considered by a broad range of users
	Time consuming when used on extensive tasks.

2.1.3 Pluralistic Walkthroughs: It is an inspection method where the representative users, usability professionals and product developers meet and walkthrough the paper-based proposed interface and scenarios step by step assuming the role of real user.The participants are provided with the hard copies of the screens/panels as it would appear to the real users while working on the final product. The participants

are asked to note down the actions in response to the designated tasks. Each dialog element and action suggested for the task is discussed in length amongst the participants with the view to identify the potential usability problems. The Pluralistic walkthrough method is prized for its ability to evaluate the design directly from the users even before a functional prototype is accessible (Riihiaho, 2002). However, due to lack of functional model, the method fails to reliably evaluate the efficiency of the system, its navigation across the interface and the user interface flow(Bias, 1994). Some of its advantages and disadvantages are (Nielsen, 1994),(Gamberin & Valentini, 2003) are:

Advantages	Disadvantages
Large number of usability issues can be identified at the same time.	Organizing and scheduling a group can be a challenging task.
Quick resolution of the usability problems.	The movement of the group depends on its slowest member.
Design can be evaluated directly by the user even when no functional prototype exists.	Difficult to evaluate the efficiency and navigational features of the product.

2.1.4 Feature Inspection: As the name suggests, this technique emphasizes on the features of the product (Bell, 1992). The inspectors are provided with sequential list of features along with situations to complete a task. Each feature is then analyzed to see whether it is available, understandable and useful. The idea is to ensure easy accessibility of the features so that the users are able to get through them without any trouble. This method can be used to evaluate the features during coding, testing, and deployment stages Some of the advantages and disadvantages of this technique (Jordan, 1998) are:

Advantages	Disadvantages
Carry out product inspection and usability evaluation at the same time.	Does not reflect user's experience
More significantly used in website evaluations	Provides an overview of the product usage without any minute details.
	Cannot measure usability.

2.1.5 Thinking aloud : The development of this testing protocol is credited to Ericsson and Simon (1984). It requires the participants to use the system and speak out their thoughts as they work through the user interface. The users are made to articulate whatever they are feeling, thinking, doing while using the system. The idea is to get an insight into users' expectations and to identify aspects that cause usability problems. The method also reveals how the user actually thinks about that product and whether it matches with what it intended to provide. Since the method slows down the performance of the user, it is not very useful in measuring time on tasks. The user may also find it difficult to work and speak out simultaneously. The user may also consciously or unconsciously miss out things as they talk(Nielsen, 2012). (Cotton, & Gretsy, 2006) deployed this method to evaluate an online biological resource for nurses. The study highlighted that it was important to guide the users about the thoughts they should pay attention to and articulate.They also found it unrealistic to expect the user to read the text on the screen and verbalize their thoughts at the same time. However, the authors reported that the richness of data collected through this method overshadows the limitations of this technique. Some of the advantages of think aloud protocol are as follows(Nielsen, 2012):

Advantages	Disadvantages
Cost effective as the method does not require special infrastructure to run the test.	Quite un-natural and difficult for the user to indulge in monologue throughout the long process.
Can be used at any stage of process lifecycle :early paper prototypes or final release.	The user may avoid giving candid statements.

Articulating thoughts can give answers to what issues arise and why they arise.	The method does not provide any quantitative data and hence restricts performance measurement

2.1.6 Perspective-based Usability Inspection: With a view to develop an inspection technique that inspects the usability from a single perspective at a time, Zhang et al.(1998) devised a perspective-based usability inspection method. According to this technique a wide range of usability issues are divided along different perspectives and each inspection session is made to lay stress on one perspective at a time. The study inspected three usability perspectives: novice use, expert use, and error handling and declared higher success rates with this method. The procedure for perspective –based usability inspection involves the following steps:

1. Defining the different perspectives for the product/service.
2. Providing background information to understand the perspective at its best.
3. Defining specific tasks, actions, features important for each perspective.
4. Providing forms to the inspection team to record their findings about usability issues.
5. Reviewing the perspectives and reporting forms.
6. Allowing inspectors to conduct the inspection independently.
7. Compiling the results of different inspectors to reach out to the total usability issues noted for the system.

The advantages and disadvantages of this technique are as follows:

Advantages	Disadvantages
Focused attention on a single perspective broadens the problem finding ability.	Some perspectives can be difficult to visualize and work on.

2.2 USABILITY TESTING METHODS

2.2.1 Remote Usability Evaluation: This usability testing allows the facilitators to test the usability of the user-interface with the participants positioned in their comfort zones (homes, offices, etc). This means that

the participants and facilitators are present at two geographically different locations. The facilitators observe participant's behaviour and capture their responses distantly through the screen-sharing software or some remote online usability tools. These tests could either be moderated, where both the participants and facilitators are online at the same time working in a real environment or un-moderated , where the users do not have any remote/local real time support. A study conducted by Tullis et al.(2002) compared the lab and remote usability testing of the websites and suggested that both types of usability testing captured similar information about the usability issues and displayed high correlation for task completion and task time data. Some of the advantages and disadvantages of this approach are as follows:

Advantages	Disadvantages
Enables administering a larger participant set as compared to what can be accommodated in a lab environment.	Testing sensitive data can compromise the security.
Accommodates a diverse group of test participants from distant locations.	Restricted view of user's body language can be a limiting factor in getting in-depth user's reaction to the content under study.

2.2.2 Coaching Method: This method (Nielsen , 1993)enables the test participants to seek answers from the experts on any system-related issue during the usability testing process. The expert, who acts as a

coach, answers the users' questions in the best possible manner and gathers their information needs with a view to provide better training and documentation support. This method also allows the experts to redesign the interface in a way that clarifies such doubts and confusions for future. Usually the evaluator/tester acts as a coach but in another variant of this method, a separate coach is chosen from amongst the expert users and this enables the evaluator to keep track of how the coach has answered a participant's query. Some of the advantages and disadvantages of this method are:

Advantages	Disadvantages
This method enables the novice learners to easily grasp the system.	The method is relatively time consuming
Involves the user in the evaluation process	The coach performs a dual role. He has to act as a coach and as an evaluator

2.2.3 Co-Discovery Learning: During this test(Nielsen, 1993) two or more test participants are asked to complete the tasks together, while the evaluator observes their interaction. The participants are expected to help each other in order to accomplish the task ,as they would do in their normal work environment. They are also motivated to verbalize their feelings, experience, impression as they work through the system. This method is an adaptation of think aloud protocol and is far more natural to use as in real situations people tend to collaborate while working and they are likely to vocalize their feelings with colleagues around. The advantages and disadvantages(Trulock) are as follows

Advantages	Disadvantages
It is more realistic as compared to a single user scenario	Aspects which are usable but not optimal (eg: poor navigational structure) are hard to diagnose through this method.
This method gives a clear picture of how the system is being used and how the users feel about it.	Users may accomplish the assigned tasks despite problems, either through determination or learning.

2.2.4 Performance Measurement:This method involves usability evaluation of a system working in laboratory simulating the real environment to determine the usability issues and to compare measures such as no. of user errors, task time , success rate, etc. This method is used to obtain quantitative data about test users' performance during the usability testing process(Nielsen, 1993). The test is performed in a usability laboratory so as to keep the distractions to the minimum. About 5-8 test users are considered desirable for obtaining reliable results. MUSiC Performance measurement model (Macleod et al., 1997) is an example of this method. Some of the quantitative measurements that can be possibly obtained through performance measurement method are:

- The time taken by users to complete a specific task.
- The number of task of various kinds that can be completed within a given time limit.
- The Ratio between successful interactions and errors.
- The time spent recovering from errors.
- The number of user errors.
- The number of commands or other features that were never used by the user.
- The number of system features the user can remember during a debriefing after the test.
- The frequency of use of the manuals and/or the help system, and the time spent using them.

This technique is usually used along with retrospective testing or questionnaires so that both quantitative and qualitative data can be obtained. Some of the advantages and disadvantages of this method include:

Advantages	Disadvantages

Quite a lot of usability problems relating to specific skills, users expectations, etc can be easily determined.	Users are required to act naturally in an unnatural setting.
Measures can be acquired for gauging users' efficiency and effectiveness.	Distractions can affect the accuracy of quantitative data collection

2.2.5 Retrospective Testing: This techniques(Nielsen, 1993) involves viewing the video recording of the test sessions by the testers along with the test participants to understand the participants' behavior during the test. The tester asks a series of questions to the participant while viewing the video to better analyze his actions. This technique is usually used in combination with other methods especially with those that restrict the interaction amongst the testers and participants during the testing session. Some of the advantages and disadvantages of this method are :

Advantages	Disadvantages
Minute details can be analyzed through the video recordings.	The technique cannot be used on its own , it needs to be used along with other methods.
The analysis of the users' action gives an in- depth insight into their thought process.	It is time consuming as the time taken by each test is at least twice as long .
	Every user's interaction with the computer needs to be captured for future analysis.

2.2.6 Question Asking Protocol: under this technique the user is provided with the system to be evaluated along with a set of tasks to be performed. While performing the tasks, he is asked to vocalize his thoughts and feelings about the product just as it happens in Think aloud protocol. Apart from this, the tester asks a lot of direct questions about the system in question to gather how the interface is being interpreted by the user. Testers may also prompt the users to comment on those aspects of the design which the user may have neglected. The users' responses to these questions help the testers to understand their mental model of the system (Dumas & Redish, 1993).

Advantages	Disadvantages
It sheds light on what problems the user encounters in what context.	Probing the participants may disturb their natural thought process
More natural to verbalize thoughts than think aloud.	The success of the method depends largely on the kind of questions raised during the course of testing. Vague questions or responses can deter it.

2.2.7 Shadowing Method: In this usability evaluation technique (Nielsen, 1993) an expert user accompanies a test participant like a shadow within his natural environment to observe his behavior as he uses the product. The expert user is a mere observer and he is not suppose to interfere with the participant's work as that can alter the way a participant behaves with respect to any situation. The expert user then explains this behavior to the tester during the testing session. This technique is an alternate to think aloud protocol and used in situations where the test participant is not suppose to think aloud or interact with the tester while performing his tasks. Some of the advantages and disadvantages of this technique are as follows(Anonymous , 2012).

Advantages	Disadvantages
All observations are made in the natural environment increasing the authenticity of the data captured.	The situation could be mis-represented or mis-understood by the observer.

The presence of observer enhances the observation as information about body language, pace, mood, motivation level can give a complete picture of users' view.	Requires a lot of effort on observer's part as a lot of accurate data needs to be generated.

2.2.8 Teaching Method: Under this technique (Vora & M. Helander, 1995) the test participant is made to interact with the system so that he gets familiar with the process of task completion. Later, the test participant is asked to mentor a naïve user about the working of the system. The naïve users are very specifically instructed to restrict their participation and be dormant when it comes to problem solving.

Advantages	Disadvantages
The number of verbalization in this method by the test participants is much more than the verbalizations using think aloud protocol method.	Since the test user himself is not an expert in the field , mentoring a naïve user is like blind leading a blind.

2.2.9 Rapid Interactive Testing and Evaluation (RITE): This method is quite similar to the traditional usability testing. According to this method, the test participants are chosen and brought to the lab, decision is taken regarding how the participant's behavior would be measured, test scripts are prepared and finally the participants are involved in a verbal protocol such as think aloud. The changes to the user interface are made rapidly i.e as soon as an issue is identified, at times it could be even after observing the very first participant(Medlock, 2002).

Advantages	Disadvantages
The changes are made quite rapidly resulting in putting forward a better model for each test participant.	Poor resolution of an issue can distort other parts of the user interface.
Helps in improving the final product quickly and efficiently.	Too many changes at once can mess up the system.

2.3 Inquiry Methods

2.3.1 Field Observations: This method(Nielsen, 1993) entails the testers to visit the user's workplace and observe them work in their natural setting. The testers make a note of how the users deal with the system and accomplish their tasks. They also try to understand the users' thought process with respect to the use of the system. The testers may even inquire about the user's job profile and ways in which he uses the system with a view to collect as much data as possible. The observation process may be Direct, where the tester is present while the task is being done or Indirect, where the task is monitored by the tester through cameras and video recordings. Direct observation allows the testers to emphasize on the areas of interest while the indirect observations aid in capturing activities that would otherwise have been overlooked (Preece et al., 1994). Some of the advantages and disadvantages of this method includes:

Advantages	Disadvantages
It is carried out in the real environment giving a better picture of how the system fits into natural setting.	This technique can be obtrusive as the users may amend their behavior and working style due to the presence of the tester.
	Arranging for field visits, observations, user interviews, data analysis, etc are expensive and time consuming processes.

2.3.2 Focus Groups: Under this evaluation technique, multiple people discuss and analyze a topic with a view to make refinements in the product. A moderator prepares a list of discussion points that the group needs to address and captures their reactions and inputs. The study conducted by Tee et al.(2013) is an

example of this method, where open ended questions were posed to the focus groups to get opinions about the interface of Moodle platform. Focus groups provide valuable feedbacks for removing the weaknesses and strengthening the final product. This was well reported by the work done by Parker et al.(2015) where the focus group recommended modifications to the software to enable better and structured understanding of the content by the students.

Advantages	Disadvantages
Useful to capture spontaneous responses from the group. Discussion may also give rise to new ideas.	The collected information is likely to have low validity and is hard to analyze due to its unstructured nature.

2.3.3 Questionnaires : Some authors are of the opinion that seeking responses directly from the stakeholders(users) about the use of the system is the most suitable way to detect usability issues.(Ardito et al., 2006; Dix et al., 2004) This approach involves the use of survey methods such as questionnaires. Questionnaire is a set of written questions asked to a selective set of respondents post the usage of the system with the view to know their satisfaction level with respect to effectiveness of the system in question. There are many questionnaires reported in the literature for evaluating usability of interfaces such as QUIS[(Chin et al., 1988), PUEU (Davis, 1989), NAU(Nielsen, 1993), CSUQ(Lewis, 1995), PUTQ(Lin et al., 1997) ,USE(Lund, 2008), PHUE-1997(Perlman, 1994), While some people have customized and used the above questionnaires to suit their requirements, others have devised new set of questionnaires specific to their domains. Holzinger(2005)pointed out that though questionnaires are a tool to statistically measure the users' preferences and opinions but only data from a large number of respondents can guarantee significance.

Advantages	Disadvantages
The Questionnaires can explore the learner-oriented issues, which are at times missed out by the experts during heuristic evaluations.	Questionnaires cannot be customized for individual users as compared to personal interview method of survey.
They are easy to deploy, quick and economical to manage and store.	There is no mechanism to check mis-interpretations of the questions by the respondents.
They can have a wider geographical reach.	
A large amount of data can be collected and outcome can be reliably compared.	
Open ended questions allow users to voice out their issues and expectations from the system.	

2.3.4 Logging actual use: This technique(Nielsen, 1993) involves having the computer software automatically record the statistics about the use of the system. This enables the evaluators to capture a lot of data pertaining to frequently used features of the system, least used features, aspects that generated errors or parts of the system that required much help, how the user moved from one screen to another etc. Ivory and Hearst(2001) are of the view that automation is a promising technique to augment the existing usability evaluation techniques. It not only helps to study the users 'comprehensive use of the system but also aids in finding the usability issues encountered by them which may otherwise get ignored.

Advantages	Disadvantages
The data collected is from the real environment	Interpreting and analyzing the logs are time consuming tasks.

Automation enables collecting data from large no. of users working in varied circumstances.	The technique cannot be used as an independent method and needs to be clubbed with other techniques as it only highlights what the users did and not why they did it that way.

2.3.5 Interviews: In this technique, the evaluators prepare a list of questions based on specific interest areas and ask these questions to the users. The responses to these questions bring a lot of detailed data. The interview can either be structured, where the specific questions are guided by special agendas or unstructured, where the data collection is not concerned with any specific area of the system but aims to gather general information about the use of the system(Nielsen,1993).

Advantages	Disadvantages
Works well for gathering information that can be obtained only through the interaction between the users and interviewers.	Every interview needs to be recorded for future analysis, which may be a costly process.
	The questions have to be very carefully framed without any biasness so that it does not obstruct the flow of thoughts.

Evaluation Method	Applicable Stages in Development Life Cycle				
	Requirement	Design	Code	Test	Deployment
Proactive Field study	∫	∫			
Pluralistic Walkthrough		∫			
Teaching Method		∫	∫	∫	∫
Shadow Method		∫	∫	∫	∫
Co-discovery Learning		∫	∫	∫	∫
Heuristic Evaluation		∫	∫	∫	∫
Think aloud Protocol		∫	∫	∫	∫
Cognitive Walkthroughs		∫	∫	∫	∫
Coaching Method		∫	∫	∫	∫
Performance Measurement		∫	∫	∫	∫
Interviews		∫	∫	∫	∫
Retrospective Testing		∫	∫	∫	∫
Remote Testing		∫	∫	∫	∫
Feature Inspection			∫	∫	∫
Focus Groups				∫	∫
Questionnaires		∫	∫	∫	∫
Field Observation				∫	∫
Logging Actual Use				∫	∫
Perspective Based		∫	∫	∫	∫
Question Asking Protocol		∫	∫	∫	∫
Rapid Interactive Testing		∫	∫	∫	

Table 1:
Summary of usability evaluation techniques and their applicable stages indevelopment life cycle

3. Usability Evaluation Techniques for E-Learning Systems: e-learning is the buzz word today. More and more people are resorting to online courses for upgrading or acquiring new skills. Just like any other system, the success of e-learning systems also depends to a great extent on its usability. The increasing diversity of learners, dynamics of the evolving technology and drastic changes in the learning tasks pose challenges to the usability evaluation of e-learning systems.

3.1 Heuristic Evaluation in the field of e-learning: HE finds extensive application in the field of e- learning(Benson et al., 2002),(Albion,1999) ,(Ardito et al., 2006),(Pipan et al. , 2006) ,(Zaharias,2007). It is one of the most cost effective methods used in this field(Albion,1999).Much work in this field involves the direct application of Nielson's heuristics or a slight variation of it (Dringus and Cohen, 2005), (Quinn,1996)(Albion,1999). For example Squires and Preece(1999), integrated the Neilson's usability aspects with the learning issues to drafts heuristics based on socio-constructivism principles. Benson et al.2002) proposed a set of fifteen heuristics that combined Nielsen's original ten heuristics (developed for software in general) with usability attributes specific to e-learning programs This hybrid heuristics as well as the protocol guided the evaluation process in Sulaiman(2002). Mehlenbacher (2005)provided a more comprehensive view of e-learning evaluation by combining design, human-computer interaction research, usability and rhetorical theory. Davids(2013) deployed six experts from the fields of usability, e-learning, instructional design, medical informatics, and the content area of nephrology and conducted heuristic evaluation for the multimedia e-learning resource followed by usability testing. They concluded that Heuristic evaluation successfully uncovered most of the serious problems but it needs to be combined with user testing for optimized results.

Some of the indicative problems of HE emerged from the review are as follows:

- The traditional Heuristics need to be intensified to deal with the specific domain.
- Evaluators subjective judgment determines the fate of the evaluation process.
- HE's effectiveness can be improved by integrating it with user testing.

3.2 Other Methods used in the Industry: While the literatures demonstrates an extensive use of existing usability evaluation techniques, some researchers came up with new methods or extended and tweaked existing ones. For example Costabile et al.(2005) described the SUE(Systematic Usability Evaluation) methodology to evaluate usability-related issues of an e-learning system. The technique is guided by evaluation patterns called abstract tasks, which detail out the activities that drive the inspector's evaluation process. A pilot study was conducted to identify significant features of e-learning interaction. A number of problems were pointed out during the interview and observation done with the help of think aloud method. The platform was evaluated with respect to

four dimensions namely: Presentation, Hypermediality, Application proactivety and user activity. Triacca et al.(2004) presented a methodical technique called MiLE for usability evaluating of e- learning platforms. This approach uses scenario-based inspection method along with heuristic- oriented evaluation to conduct a learner –centric validation process. It takes into account factors such user profile, their goals, usability attributes, scenarios of use and provides guidelines to content developers and designers for a practical and cost effective usability evaluation. The approach is a proven methodology to foresee usability obstacles and to give suggestions for efficient redesigns.In order to determine the change in user's usability perception with the improvement in a particular usability problem, Oztekin et al.(2013) devised a machine learning based e-learning usability evaluation technique. Models based on multiple linear regression, decision trees, support vector machines and neural networks are used to design prediction models that establish the relationship between the system usability and its predictor factors. Senstivity and pareto-like analysis are conducted to find out the ordering of the predictors on the basis of their importance. The proposed methodology claims to identify the most significant usability indicators.To identify the user's perception on the usability of a moodle platform interface, Tee et al.(2013) carried out a user testing evaluation method. The users' opinion was gathered using focus group interviews, which was followed by a brainstorming session that helped to determine the solutions to improve the usability issues in the moodle applications.

Oztekin et al.(2010) proposed a quantitative usability evaluation method based on a novel checklist called UseLearn. UseLearn consists of quality and usability dimensions and sets up a strong affinity between e-learning quality and usability. The approach uses structural equation modeling and criticality metric analysis to identify the most significant usability problems. The checklist along with the quantitative methodology provides an effective guideline to emphasize on pertinent usability problems to improve the overall usability of the system. Wabil and Khalifa (2009) developed a framework for combining varied usability evaluation methods to effectively evaluate the usability of an e-learning portal. The framework is build with the objective of harmonizing the potentials and weaknesses of the usability evaluation method with the categorization of usability problems. Availability of usability evaluation alternatives, goal of the application and an understanding of the strengths and weaknesses of each evaluation method influence the choice of the usability evaluation methods to be integrated in the framework. The given framework demonstrated a wider coverage of usability issues during its application in evaluating the usability of Mawhiba web portal.

4. Conclusion: There are various usability evaluation methods that exist in the literature. Each one of them has its own share of advantages and disadvantages. No one single technique can be applied satisfactorily on all types of situations as the usability evaluation depends greatly on the product features, the user characteristics, their goals and the context of use. Literature is full of examples where one or more methods have been tweaked or combined to evaluate the usability of an e- learning system depending on determinants such as type of system, suitability of the method, resources and time available, type of users, concerned usability factors, etc.

Considering that usability is one of the core determinants of e-learning's success, it is critical to come up with an evaluation mechanism that impacts the "real" e-learning design and development. There is a need to have a new framework that helps to assess the usability of an e-learning portal in a more holistic way so that more promising learner-centric solutions can be created. Also the fact that various new measures of usability (as apparent in the literature review) are emerging, this in itself

calls for revisiting the concept and suggesting an integrated approach to usability evaluation. In the near future, we intend to propose one such model.

References

Alsumait, A., and Al-Osaimi, A.,2009 ,Usability heuristics evaluation for child e-learning applications. In Proceedings of the 11th International Conference on Information Integration and Web-based Applications & Services, iiWAS '09, ACM, pp. 425—430

Collazos,C.A., Ortega,M.,Granollers,A., Luna,H. and Fardoun,H., 2017 "Usability Evaluation Trends in Ibero-American Countries," in IT Professional, vol. 19, no. 5, pp. 61-64, doi:10.1109/MITP.2017.3680965

Gerhardt-Powals, J., 1996. Cognitive engineering principles for enhancing human-computer performance. International Journal of Human-Computer Interaction 8, 189–211

Tognazzini, B., 2003. First Principle of Interaction Design . available on :www.asktog.com/atc/principles-of-interaction-design/ accessed on 19th nov 2017.

Schneiderman, B.1998. Designing the user Interface: Strategies for effective human-computer interaction. Addison Wesley Longman, Reading, MA.

Desurvire H, Caplan M, Toth JA, 2004, Using heuristics to evaluate the playability of games. Proceedings of the CHI 2004 Conference on Human Factors in Computing Systems, April 24–29,Vienna, Austria.

Rau P, Liang SF, 2003, Internationalization and localization: evaluating and testing a website for Asian users. Ergonomics, 46, 1-3, 255-270.

Nielsen, J., 1994, Heuristic Evaluations. In: J. Nielsen and R.L. Mack. (Eds), Usability Inspection Methods. New York: John Wiley & Sons.

Karoulis, A. and Pombortsis, A., 2003, Heuristic Evaluation of Web-Based ODL Programs. In: C. Ghaoui.(Ed.), Usability Evaluation of Online Learning Programs.Hershey, P.A.: Information Science Publishing.

Jeffries, R., Miller, J. R., Wharton, C., and Uyeda, K. M., 1991, User interface evaluation in the real world: A comparison of four techniques, In CHI '91 Conference Proceedings (pp.119-124), New York: ACM Press.

Karat, C., Campbell, R. L., and Fiegel, T.,1992, Comparison of empirical testing and walkthrough methods in user interface evaluation, In CHI '92 Conference Proceedings (pp. 397-404), New York: ACM Press.

Desurvire, H. W., and Thomas, J. C., 1993, Enhancing the performance of interface evaluators. In Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting (pp. 1132-1136), Santa Monica, CA: Human Factors and Ergonomics Society.

Virzi, R., Sorce, J., and Herbert, L. B., 1993, A comparison of three usability evaluation methods: Heuristic, think-aloud, and performance testing,In Proceedings of the Human Factors and Ergonomics Society 36th Annual Meeting , pp. 309-313, Santa Monica, CA: Human Factors and Ergonomics Society.

Lund, A. M., 1998, The need for a standardized set of usability metrics, In Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting, pp. 688-691,Santa Monica, CA: Human Factors and Ergonomics Society.

Van den Haak, M.J., De Jong, M.D.T., Schellens, P.J., 2004, Employing think-aloud protocols and constructive interaction to test the usability of online library catalogues: a methodological comparison. Interacting with Computers 16, 1153–1170

Doubleday, A., Ryan, M., Springett, M., Sutcliffe, A., 1997, A comparison of usability techniques for evaluating design, In: Proceedings of Designing interactive systems.

Molich, R., Ede, M.R., Kaasgaard, K., Karyukin, B., 2004, Comparative usability evaluation. Behaviour and Information Technology 23(1), 65–74.

Hertzum, M., Jacobsen, N.E., 2001, The Evaluator Effect: A Chilling Fact about Usability Evaluation Methods. International Journal of Human-Computer Interaction 13(4), 421–443.

Lund, A., 1997,"Expert Ratings of Usability Maxims. *Ergonomics in Design*, Volume 3, Issue 5.

Macefield, R., 2014, An Overview of Expert Heuristic Evaluations, https://www.uxmatters.com/mt/archives/2014/06/an-overview-of-expert-heuristic-evaluations.php accessed on 23 Nov, 2017

Oztekin ,A., Delen,D. , Turkyilmaz A. and Zaim,S., 2013, "A machine learning-based usability evaluation method for eLearning systems", Decision Support Systems 56 , pp 63–73.

Benson, L., Elliott, D., Grant, M., Holschuh, D., Kim, B., Kim, H., Lauber, E., Loh, S. and Reeves, T.C. ,2002, Usability and Instructional Design Heuristics for E-Learning Evaluation. In P. Barker & S. Rebelsky (Eds.), *Proceedings of ED-MEDIA 2002--World Conference on Educational Multimedia, Hypermedia & Telecommunications* (pp. 1615-1621). Denver, Colorado, USA: Association for the Advancement of Computing in Education (AACE).

Sulaiman, J., Zulkifli, T., Ibrahim, K.S.K., Mohd Noor, N.K.,2009, Implementing Usability Attributes In E-learning System Using Hybrid Heuristics. Published in: Information and Multimedia Technology, ICIMT '09

Alfadhli,S., Alsumait, A.,2015,Game-Based Learning Guidelines: Designing for Learning and Fun , 2015 International Conference on Computational Science and Computational Intelligence, IEEE

Davids M.R., Chikte, U.M., Halperin M.L., 2013, An efficient approach to improve the usability of e-learning resources: the role of heuristic evaluation, , 2013 Sep,37(3):242-8. doi: 10.1152/advan.00043.2013.

Albion, P.R.,1999, Heuristic evaluation of educational multimedia: from theory to practice, Proceedings of 16th Annual Conference of the Australasian Society for Computers in learning in Tertiary Education, ASCILITE.

Ardito, C., Costabile, F., De Marsico, M., Lanzilotti, R., Levialdi, S., Roselli, Z.T. and Rossano, V., 2006, "An approach to usability evaluation of e-learning applications", Univ Access Inf Soc., Vol. 4, pp. 270-83.

Pipan, M., Arh, T. and Jerman-Blaz ˇic ˇ, B, 2006, The evaluation cycle management – model for the assessment of the usability

and applicability of learning management systems, Is information technology shaping the future of higher education?, Proceedings of the 12th International Conference of European University Information Systems, EUNIS 2006, Estonia, pp. 325-32.

Zaharias, P.,2007, Heuristic evaluation in e-learning context: selecting the appropriate tasks and reporting usability problems, Proceedings of International Conference on E-Learning– ICEL07.

Dringus, L.P. and Cohen, M.S., 2005, An adaptable usability heuristic checklist for onlinecourses, Proceedings of Frontiers in Education, 35th Annual Conference.

Quinn, C.N.,1996, Pragmatic evaluation: lessons from usability, Proceedings of 13th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education.

Squires, D. and Preece, J., 1999, Predicting quality in educational software: evaluating for learning, usability and the synergy between them, Interacting with Computers, Vol. 11 No. 5, pp. 467-83.

Mehlenbacher, B., Bennett, L., Bird, T., Ivey, M., Lucas, J., Morton, J. and Whitman, L. ,2005, Usable e-learning: a conceptual model for evaluation and design, Proceedings of HCI International 2005: 11th International Conference on Human-Computer Interaction,Volume 4 – Theories, Models, and Processes in HCI, Mira Digital, Las Vegas, NV, pp. 1-10.

Chin, John P. , Diehl, Virginia A. , 1988, Development of an Instrument Measuring User Satisfaction of the Human-Computer Interface *Interface Evaluations* / *Proceedings of ACM CHI'88* Conference on Human Factors in Computing Systems 1988-05- 15 p.213-218

Davis, F. D., 1989, Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology ,MIS Quarterly , v.13 n.3 p.319-340

Nielsen, J. ,1993, Usability Engineering, Academic Press.

Lewis, J. R.,1995, IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. International Journal of Human-Computer Interaction, 7:1, 57-78.

Lin, Han X. Choong, Yee-Yin, Salvendy, Gavriel, 1997, "A Proposed Index of Usability: A Method for Comparing the Relative Usability of Different Software Systems" *Usability Evaluation Methods* Behaviour and Information Technology, v.16 n.4/5 p.267-278.

Lund, Arnold (Arnie) M, 2008, Measuring Usability with the USE Questionnaire, 2008-09-05.

Tee, S. S., Wook, T. S. M. T., Zainudin, S., 2013. User testing for Moodle application. International Journal of Software Engineering and Its Applications , 7(5), pp. 243-252.

Parker, S., Mayner, L., and Michael Gillham, D.,2015, E-learning for Critical Thinking: Using Nominal Focus Group Method to Inform Software Content and Design. *Nursing and Midwifery Studies*, *4*(4), e30471. http://doi.org/10.17795/nmsjournal30471

Holzinger, A., 2005 Usability engineering methods for software developers. Communications of the ACM , Vol. 48, No. 1 (Jan. 2005), pp. 71-74

Perlman, G., 1995, Practical usability evaluation, Proc. CHI 1996 Conf. Human Factors in Computing Systems, Pages 407-408 Dix, A., Finlay, J., Abowd, G., Beale, R., 2004, *Human-Computer Interaction*,(Third Edition), Harlow Assex: Pearson Education Limited.

Riihiaho,S., 2002, The Pluralistic Usability Walk-Through Method, Ergonomics in Design: The Quarterly of Human Factors App., vol. 10, pp. 23-27.

BIAS, R.,1994, The pluralistic usability walkthrough: Coordinated empathies, In Usability inspection methods, NIELSEN, J. and MACK, R.L. (eds.), New York, Wiley. pp. 63-76.

J. Rubin Handbook of usability testing: how to plan, design and conduct effective tests,1994, Wiley, New York

Nielsen, J.,2012, Thinking Aloud: The #1 Usability Tool, available at https://www.nngroup.com/articles/thinking-aloud-the-1-usability-tool/

Yen, P. Y., Bakken, S., 2009, A Comparison of Usability Evaluation Methods: Heuristic Evaluation versus End-User Think-Aloud Protocol – An Example from a Web-based Communication Tool for Nurse Scheduling, in AMIA ... Annual Symposium

proceedings / AMIA Symposium, AMIA Symposium 2009:714-8 ·

Bell, B. ,1992, Using programming walkthroughs to design a visual language, Technical Report CU-CS-581-92 (Ph.D. Thesis), University of Colorado, Boulder, CO.

Jordan, P.W., 1998, An introduction to Usability, Taylor & Francis Ltd.

Gamberini,L.,Valentini E., 2003, Web Usability Today: Theories, Approach and Methods, IOS Press. Nielsen,J., 1994, Usability

Inspection Methods, John Wiley and Sons, Inc.

Rowley,D., Rhoades D.,1992, The cognitive jogthrough: a fast-paced user interface evaluation procedure, In Proceedings of the ACM CHI'92 Conference on Human Factors in Computing Systems, ACM Press, pp. 389-395.

Green,T., Burnett,M., 2000, Using the Cognitive Walkthrough to Improve the Design of a Visual Programming Experiment, IEEE.

Zhang, Z., Basili,V., and Shneiderman, B., 1998, An empirical study of perspective-based usability inspection, by *Proceedings of the Human Factors and Ergonomics Soceity 42nd Annual Meeting*. pp. 1346-1350, Chicago.

Tullis,T., Fleischman,S., McNulty,M., Cianchette,C., and Bergel,M., 2002, An Empirical Comparison of Lab and Remote Usability Testing of Web Sites, Presented at the Usability Professionals Association Conference, July 2002, Orlando, FL.

Trulock,V., " Understanding HCI", http://hci.ilikecake.ie/eval_codiscoverylearning.htm accesssed on 20/7/2018
Macleod, M., Bowden, R., Bevan, N. and Curson, I., 1997, The MUSiC Performance Measurement Method. Behaviour and Information Technology, 16.

Dumas J. S. and Redish J. C., 1993, A Practical Guide to Usability Testing.

Anon ymous , 2012, Focusing on Methods: Shadowing , avialble at https://experience.sap.com/basics/post-27/ accessed on 15/07/2018

Vora, P. and Helander, M., 1995, A Teaching method as an alternative to the concurrent think-aloud method for usablity testing, in Y. Anzai, K. Ogawa & H. Mori "Symbiosis of Human and Artifact", pp.375-380.

Medlock M. C., Wixon D., Terrano M., Romero R. and Fulton B.,2002, Using the RITE method to improve products: A definition and a case study, Presented at the Usability Professionsals Association 2002, Orlando Florida.

Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S. and Carey, T., 1994, Human-Computer Interaction. Reading MA: Addison-Wesley.

Ivory, M.Y. and Hearst, M.A, 2001, The state of the art in automating usability evaluation of user interfaces, ACM Computing Surveys 33: 470–516.

Cotton, D., & Gretsy, K., 2006, Reflecting on the think-aloud method for evaluating e-learning. British Journal of Educational Technology , 37 (1), 54–54.

Matera,M., Costabile,M.F., Garzotto,F. and Paolini,P., 2002, SUE Inspection: an Effective Method for Systematic Usability Evaluation of Hypermedia, IEEE Transactions on Systems, Man and Cybernetics- Part A, Vol. 32, N. 1, 93-103.
De Angeli, Matera,M., Costabile,M.F., Garzotto,F., and Paolini,P., 2000, Evaluating the SUE Inspection Technique, Proc. AVI'2000 – International Conference on Advanced Visual Interfaces, Palermo, Italy, May 24-27, ACM Press, 143-150. Triacca, L., Bolchini, D., Botturi, L., & Inversini, A., 2004. Mile: Systematic usability evaluation for e-Learning web applications.AACE Journal ,12(4).

Zaharias, P., and Koutsabasis, P., 2011 , Heuristic Evaluation of e-Learning: comparing two Heuristic Sets. Journal of Campus-Wide Information Systems, Vol. 29, Issue: 1, pp.45 - 60.