Development of Landscape of Usability Evaluation Methods for Mobile Applications

Sunday Ariyo Ajibola
School of Computing
University of South Africa
Pretoria, South Africa
ajibolasunny@gmail.com

Leila Goosen
Department for Science and Technology Education
University of South Africa
Pretoria, South Africa

Abstract—*The sales of mobile technologies have grown substantially as well as the significant increment in the development of mobile applications. Mobile devices like smartphones and tablets, among others, have gained popularity and enhanced users' interactions and the exchange of information. The major aim of this study is to conduct extensive mapping of related literature sources regarding the usability methods in the context of mobile applications. The study seeks to uncover the most used set of usability evaluation methods in the context of mobile applications and to answer the question of when to use which usability evaluation methods. In order to have a proper understanding of when to use which method, this study presents the landscape of usability evaluation methods and categorizes them according to 3-dimensional approaches with axes. The results prove that different usability evaluation methods have been used for the evaluation of mobile applications. In addition, the study uncovered 19 usability evaluation methods for the evaluation of mobile applications. The findings show that Survey, In-lab studies, Interview, Observation, and Log file are the most widely used methods, which account for more than 10 percent each of the reviewed studies*

Keywords- Usability evaluation methods; usability methods; mobile usability issues; mobile usability; mobile usability methods; mobile e-commerce usability

I. INTRODUCTION

It is a widely accepted concept that the usability of a website is one of the most important requirements in website design and development. It is especially important for some specific websites, such as mobile and M-Commerce websites.

The fundamentals of website or applications' usability can be described as: It should provide all necessary information and functions that users need clearly and quickly when they are needed. In contrast, a poorly designed website with less usability will cost the owner or company with considerable operating and maintenance cost due to various negative impacts of the website [2]. In the nowadays highly competitive information age, the poor usability on commercial websites or mobile applications will only result in serious consequences with no benefits [9].

Therefore, in order to improve the usability of mobile applications, usability evaluation is necessary [7]. In the last few decades, there are different usability evaluation methods (UEMs) in order to assess the usability level in a software system or mobile applications [31]. In addition, the use of usability evaluation methods for testing the usability standard of software applications among others is increasing and becoming a standard approach in the software development processes. Hence, incorporating usability evaluation to the website development process becomes critical and necessary in order to improve the ease of use of the website or mobile applications and consequently to enhance users' experiences [44].

To the best of the researchers' knowledge, this study will have greater contribution in which usability evaluation methods are placed in a landscape in order to help usability professionals on when to use which evaluation methods in the context of mobile applications. Therefore, the main objective of this study is to uncover when to use which usability evaluation methods in the context of mobile applications. The research question is: What is the landscape

of usability evaluation methods with respect to when to use which evaluation methods in the context of mobile applications? This paper discusses the detail about usability evaluation methods in general and for mobile applications in particular and presents clearly the landscape of usability evaluation methods. In addition, based on the literature reviews of the selected 188 mobile empirical usability studies within 2005 and 2018; this research work presents the pattern, trend, and frequencies of usability evaluation methods in the context of mobile applications. The landscape presented would serve as huge benefits for usability professionals and the researchers consider it as a major contribution of this research work. To the best of the researchers' knowledge, there is a huge literature gap in this genre of study, which this study seeks to address. The next section discusses the literature review of this study.

II.   LITERATURE REVIEW

2.1   Usability Evaluation Methods

A usability study is described as a discipline that cut across many fields: Human Interface Design, Engineering, Information Architecture, Technical Communication among others [37]. Therefore, this assertion shows that usability is homeless- or has many roots.

Usability evaluation provides organized methods for attaining enhanced usability in the design of user interface during product development. The evaluation is deemed to fail except due usability engineering practice is adequately considered. Usability engineering consists of three stages: analysis of requirements, development/design/testing, and installation. The purposes of usability are achieved during the requirement stage. During the development/design/testing stage, iterative testing is done. Users are allowed to offer feedback in order to check the functionality and usability of the product at the installation stage. Getting familiar with the concept of usability evaluation is essential to understand the Usability Evaluation Methods (UEMs). Usability evaluation can be defined as an assessment of a specific software user interface or interaction method or device's functionality and usability [43].

In addition, UEMs are well-composed guidelines of activities and procedures used to gather data about the degree of the usability of the software system [67]. Generally, the usability evaluation can be divided into two parts, formative evaluation and summative evaluation [28]. The formative evaluation takes place through all the steps, from development to advanced stages of a system designing, while summative evaluation judge and assess the system design when all the steps of system design are completed. UEMs also include other methods as long as they are used for system usability evaluation in other specific applications. The outputs of UEMs are varied in different methods; some of them provide problem reports or list of usability problem as causative features or alternative solutions for current problems [6, 67].

TABLE 1. Highlights and Development Stages of Usability Evaluation Methods (adapted from [69])

Year	Usability Evaluation Methods
1970-1979	User problem documented, Platform Style Guide, Usability Labs, Metaphor [68]
1980-1989	Metrics for user performance, Rapid prototyping, UI Standard, GOMS model, Comprehensive Guidelines, Wizard of Oz, Iterative Design, Motif style guide, SUMI QUIS [69]
1990-1999	Heuristics Evaluation, Cognitive Walkthrough, Remote usability testing [28, 67]
2000+	CIF Report Format –ANSI Standard [11, 28]

Different UEMs have been developed in the last four decades in order to overcome the usability problems in software systems. Table 1 shows the general highlight and development stages of usability evaluation methods. It can be observed from Table 1 that the root of the usability evaluation methods is dated back as at the 1970s with the majority of methods are developed between 1980-1989. This may be connected to Xerox Star, which introduced the most important component of the internet – the graphical interface in 1981 and this improved and increased the reputation of the World Wide Web [42].

For the purpose of this study, this study presents and discusses only the usability evaluation methods as identified in the review of 188 mobile empirical usability studies between 2005 and 2018. The next section discusses the general classifications of usability evaluation methods.

2.2   Classifications of Usability Evaluation Methods

Usability evaluation methods can be classified based on the source utilized for the evaluation. The sources are users, models or usability experts [81]. Table 1 shows the range of historical dates for each of the usability

evaluation methods. Users were the first source used for usability evaluation to aid usability feedback, but models have attracted usability professionals for over 20 years. In addition, since the early 90s, expert-based sources like heuristics and cognitive walkthroughs have been used to gather some usability reports on software users' interfaces. However, usability professionals or usability engineers use the three methods to design, conduct and analyse the usability evaluations of the software system [21, 79].

2.2.1 User-Based Usability Evaluation Methods

Usability testing was the main evaluation method used in the 1980s and is currently still a popular method of evaluation, especially for the later part of the software design and development stages [1]. Usability testing methods can be considered as usability appraisal tactics that implement experimental testing of the interface design with typical users of such interfaces [64]. The testing process encompasses similar stages for all methods, which are identified as usability assessment methodologies. The testing process implements a test design (comprising objectives, explanations, required resources, and planning), users' choices (referring to the profiles distinct from the preceding phase), test case formulation, test case execution and, lastly, the analysis of results/outcomes [64]. In addition, prior studies show that the most extensively recognized usability evaluation approaches are user-based testing methods and expert-based methods [53, 73].

In-lab testing as an example of the user-based method is a usability testing method, which is conducted in a laboratory environment and is primarily created for usability evaluation of desktop-based systems. This testing method is deemed to encounter difficulty in mobile applications [70]. This method has different limitations in a real mobile environment as they do not represent the mobile context of use and do not have sufficient procedural control [65]. However, remote usability testing methods can be defined as a form of usability evaluation of websites or applications in which test participants and usability evaluators are detached in geographical location and/or time. The term "remote" here means the distance in location or time separation of the test participants from the evaluator(s) [11]. It can be used to conduct usability testing of user behaviour remotely and to evaluate interactions in real and natural user environments. Nowadays, many UEMs have included usability testing methods, such as questionnaire survey, in-lab test, interview, observation, log file/device data, think aloud, focus group, video/ sound/ screen recording, prototyping. Others are eye tracking, device sensor, Wizard of Oz, dairy/ camera studies, remote asynchronous testing, Web Analytics, remote synchronous testing and card sorting among others.

2.2.2 Expert-Based Usability Evaluation Methods

Usability inspection methods, also known as expert-based evaluations, refer to the method whereby the usability aspect of a user interface is monitored and investigated by one or a group of usability experts. These experts or evaluators test the user interface, locate the usability issues, and provide suggestions for improving them. Usability inspection methods are heuristic evaluation, cognitive walkthrough, and action analysis [17, 28]. Other examples of expert-based methods are feature inspection, formal usability inspection, standard inspection, consistency inspection, pluralistic walkthrough, and guideline reviews.

2.2.3 Model-Based Evaluation Methods

Compared to the above two methods, the model-based methods in usability evaluation are not commonly used. This method works based on the psychological prediction of a specific user's performance on a certain software interface. The main objective of this method is to evaluate the system usability by assessing the total task time or difficulty level of learning the task sequence of a system. An example of model-based methods are the Goals, Operators, Methods and Selection rules (GOMS) model, which perform system usability by forecasting interface functionality and time consumption on specific tasks [11, 28, 35]. The Keystroke Level Model (KLM) gives the predictions of a user's performance in a numerical format through the aid of detailed empirical studies [38]. Other examples are the Executive-Process/Interactive Control (EPIC) and Adaptive Control of Thought-Rational (ACT-R) models. The next section discusses the major challenges concerning usability evaluation methods.

2.3 Major Challenges Concerning Usability Evaluation Methods

Although, HCI experts can conduct usability evaluations to assess the usability of a product or software, the main concern in usability evaluations is to find an optimum evaluation method. The available usability methods can provide a wide range of system usability evaluations and procedures, however, each method has its own limitations and comparisons between these methods are complex and non-conclusive [28].

Briefly, these issues are presented in the following areas. Firstly, is a case where experimental usability testing techniques are employed, in order to answer general questions on the usability of the system, as an alternative to using more narrow and direct questions with commonly used experimental methods. Secondly, what measurement [35] should be used for comparison purpose? Thirdly, how to judge the result of evaluation methods on their

reliability? Each evaluation method is designed and used differently, besides, there are many shortcomings associated with their conclusions. However, there is no doubt that implementing some usability evaluation for a system before it is put into use is more beneficial than it is to take the risk of having potential usability issues without testing the product.

A commonly used practical method of implementing usability evaluation is to apply several evaluation methods in order to obtain more reliable information about the usability problem of the product [59]. Current advantages of technology in mobile and global computing can bring new opportunities for making improvements on existing usability evaluation techniques. Therefore, usability evaluation methods, which are created to evaluate the usability levels of desktop systems, such as traditional in-lab assessments will encounter difficulty in mobile applications [70]. Mobile and multi-user systems should be assessed for confidentiality and any usability matters pertaining to setting up, forming, and utilizing such policies.

The website and mobile applications creation processes go on very quickly and their completion has very short time durations. Consequently, the evaluation of usability is generally omitted in most cases in order to shorten the product development time [3]. However, usability evaluation methods provide the developer the opportunity to simulate the context of use of websites or applications by using suitable evaluation methods. Innovative evaluation methods for system usability are expected to be introduced in order to accomplish all technology-oriented tasks and the context of use of mobile applications. Future usability challenges should draw the attention of usability professionals to the need to develop suitable usability methods relating to the context of use of the applications.

Because of the peculiar nature of mobile applications, the conventional usability evaluation methods that are suitable for a laboratory environment cannot be applicable to the complex and rich nature of the natural environment of mobile applications [50]. For example, field studies and laboratory experiments have different limitations such as lack of procedural control and failure to represent the mobile context of use. Hence, there is an urgent need to present when to use which usability evaluation methods in the context of mobile applications in which this study seeks to address. The next section presents the method used in the development of the proposed landscape of usability evaluation methods for mobile applications.

III   METHOD

3.1   Techniques Used in the Review of Research Resources

In order to show the appropriateness of the proposed landscape of usability evaluation methods in the context of mobile applications, the researchers conducted a literature review on the selected resources. This study examined the general list of usability evaluation methods that are a part of usability evaluation presented in different published research studies between the year 2005 and 2018. The resources reviewed in this study were obtained from academic and non-academic sources as indicated by the literature references [33]. The researchers explored different sources for relevant literature by using keywords, sources like Google Scholar search engine, the Association for Computing Machinery (ACM) database, the Institute of Electrical and Electronics Engineers (IEEE) database, the University of South Africa (UNISA) subject databases, and other Human Computer Interaction (HCI) database sources were accessed.

Among the keywords used to search for relevant articles are: " usability methods," "e-commerce usability issues," "e-commerce usability," "e-commerce usability methods," "mobile usability issues," "mobile usability," "mobile usability methods," "mobile e-commerce usability issues," "mobile e-commerce usability," "mobile e-commerce usability methods," "mobile e-commerce usability," "e-commerce usability," "usability theory," "usability engineering," "usability studies" and "usability evaluation method".

The resources selected for this study are based on certain selection criteria that formed the basis for the inclusion and exclusion of research resources. Research resources included are those published between 2005 and 2018. The selection criteria are contingent upon whether the research resource:

1. performed an evaluation of mobile applications

2. contained software components (e.g. paper prototype) which allow users to interact with it

3. focused on users' interactions with the applications or devices and conducted an evaluation.

The method suggested by prior study is in line with the viewpoint that electronic searches only yield 10 percent of the total research resources required for a literature review [66]. The residual 90 percent were identified by going through the reference list of the research resources that had already been retrieved. The researchers determined which among them were relevant by using the inclusion selection criteria outlined above. This process is repeated until there were no relevant research resources found. The next subsection discusses in detail each of the identified usability methods as presented in the proposed landscape in the context of mobile applications.

A total of initial 1476 research papers found were retrieved by using the search strings and reading through titles of the research papers. Total of 851 of research papers were selected after reading their abstracts. The researcher then applied the three selection criteria as stated above, which reduced the total to 128 research papers that passed the selection criteria. The researcher adopted the strategy as proposed by Randolph [66] by searching through the reference lists of the already retrieved research papers that passed the selection criteria. Through this approach, a total of additional 60 research papers were discovered that passed all the selection criteria. The researcher repeated this process until there were no relevant research resources found. Therefore, a total of 188 research papers passed the selection criteria and were used in the literature review of the current research. The next section provides a brief background on e-Commerce and m-Commerce applications.

3.2   Descriptions of the Mobile Usability Methods in the Proposed Landscape

Base on the 188 relevant and selected mobile empirical usability studies, this research work identifies 19 usability evaluation methods in the context of mobile applications. Therefore, this section discusses the 19 mobile usability evaluation methods before they were placed in the proposed landscape.

1. ***Questionnaire Survey:*** This method is one of the most used methods. It is usually administered to the test participants at the end of the test in order to gain insight about their interaction with the product. A form of questionnaire is the System Usability Scale (SUS) administer to get the users' feelings about the usage of a particular product. It is usually used to get the degree of users satisfaction about the product [14]. Another example is the National Aeronautics and Space Administration Task Load Index (NASA-TLX), which is a form of a questionnaire used to measure the cognitive load of test participants [45, 48]. Intercept survey is a type of survey that is triggered while the user is interacting with live application or site while email surveys are methods that are employed by recruiting the test participants through email messages.

2. ***In-lab test***: Laboratory testing is deliberated as a customary method, which is extensively used for assessing software programs prior to their release [47]. In addition, laboratory tests on informational websites can sometimes mistakenly ignore elusive properties of writing style, navigation, and the graphical outline of the application on users' insights and presentation [19]. In addition, the method simulates users' day to day scenarios for the evaluation of the mobile device. This method of evaluation cannot adequately fit for some uncontrollable factors that have a greater impact on the real-life situation of a mobile phone. This method has different limitations in a real mobile environment, like not representing the mobile context of use and insufficient procedural control. [65]

3. ***Interview:*** In order to get qualitative data about the product under usability evaluation, interview method is usually used. This method gives insight into the reason behind users' action while interacting with the product or application. It helps the researcher to know and observe the reactions of the users in their natural context of use and assist in extracting the problem areas during the investigation [22]. For example, this method is good in conducting a usability evaluation of product prototype in order to eliminate the problem found in the existing application [26]. In most cases, this method is used in combination with other methods like field and laboratory testing method in order to get a deeper understanding of users' behaviour and reaction about the product.

4. ***Observation method***: This method involves the actual users interacting with the developed system or application [78]. This could be established through various methods, ranging from the casual observation of a particular user to complete investigations with suitable principles and control variables in order to identify the problem areas [1, 63]. In this context, observations are usually conducted in the actual implementation site or in the company's usability workrooms. Observation methods involve different analysis, like recording the observation on the observation notes, computing task completion time, recording the success and difficulty paths while completing the task and data interpretation among others [75].

5. ***Log file / Device data:*** This is an electronic record system of user's activities or interactions on a particular website or application that automatically update itself upon when each request (hit) is made through the electronic device [56]. In addition, the qualitative analysis of the log file in most cases result in the identification of usage problem of the system [56]. The auto logging nature of this method allows the collection of the visited URL history, input speed, error rate and task completion time, which can be quantitatively analysed. It can reveal if the task completion paths are in the correct order and are designed correctively. It is usually used in combination with a questionnaire or/and interview method(s) because this method is poor in the collection of qualitative data necessary to solve or address the identified usability problem [8, 77]

6. ***Heuristics Evaluation:*** This method is developed and modified by Jakob Nielsen in 1994 and is one of the most widely used usability evaluation methods [60, 63]. It comprises of a set of guidelines or heuristics use in the evaluation of user interface of the product by usability experts. It is usually conducted by 3 to 5 experts and is considered as cheap and effective compared to other usability evaluation methods,

especially laboratory method. In addition, the evaluation approaches should consider the interface design with the comprehensive assistance of heuristic techniques, which are usually based on analysis rather than on experience. These inspections are often held by industry professionals in the area of usability, whose focus is to identify prevalent design problems in the context of user interaction [64].

7. *Think Aloud:* This is one of the most widely used methods that is mostly applied in the laboratory during the usability test. The test participant is requested to articulate his feelings and thoughts as it comes to his mind or involving the test participant to loudly talk out their thoughts while interacting with the system [44, 59]. This approach gives a prompt response on the interpretation and motivations of the user in relation to the tasks and user interface elements. However, this approach of thinking aloud while performing task leads to the longer execution time of the tasks as well as limiting the user attention on the completion of the task, because of the added cognitive load- verbal effort [52].

8. *Focus group:* This is a qualitative research method, which composes of a group of selected people that gather in a place in order to ask them about their beliefs, opinions, attitudes, and perceptions in relation to a product, concept, service or idea. The participants have the freedom to express themselves with other group members and this approach is conducted in an interactive environment. In the field of usability engineering, this method is regarded as a survey method used in the collection of users' view of the software or website [10, 23]. This method is different from other ergonomics' methods because it involves several users or future users of software or application as participants in the discussion. This method usually composed of a different number of participants, which are mostly used to collect subjective data about the application usage in the usability testing [34, 71].

9. *Video/ Sound/ Screen Recording:* These methods help the evaluators to gather qualitative data with an in-depth understanding of the characteristics of the user interaction and context of use. This lead to a greater improvement of the evaluator assessments on the type, number, and nature of the identified problem area of the system [25]. In most cases, these methods are combined with other methods during the usability evaluation, like with cognitive walkthrough in order to support and to exploit data richness in relation to user contextual details and characteristics [24]. This approach helps to improve the sensitivity of the experts in the cognitive walkthrough in usage patterns and context of use during the evaluations.

10. *Prototyping:* This a method where the participants are presented with design element or developed materials in order to get their actual experience on what is most important to them and why. In this endeavour, blueprints of the software systems are produced in order to provide validation to developed functionalities. There are different modes in the prototype's perceptions, which can be haptic, visual, tactile paper and computer-based prototyping [54]. However, the most widely used is visual paper prototyping, which is mostly used to do design concepts evaluation in the product design at an early stage [54]. In addition, paper prototyping is a method used in usability evaluation in order to identify usability problem of the user interface at an early stage of product development but not suitable for the evaluation of the user interface portable devices [31, 51]. For example, the research carried out in Negara shows that experts comments help in product prototyping with the acceptable index result of the content, which gives detail information about Virtual Museum in Negara and leads to increment in the number of tourists visitation [76].

11. *Eye Tracking:* This is an electronic device that is designed to measure precisely where the test participants look on the user interface of the application while performing the task or naturally interacting with the applications or website(s). This method gives more information about user's focused points or areas during the usability test. The quality of the data is base on the device adjustment to external influence and users [52]. This method is gaining popularity in the field of usability engineering and user experience as a useful and common input medium. For example, the eye-tracking device helped in providing stable, fast and reliable means of communication for the physically challenged user who cannot use the standard keyboard and mouse [30].

12. *Device Sensor:* This method helps in limiting user involvement in gathering useful information about user activities or interaction with the system. It helps the evaluator to gather mobile sensor data, which significantly limit the involvement of the user [55]. In some cases, multiple data can be collected from different sensors in order to do a comparison of their capability on cognitive load assessment [27]. Nowadays, Smartphones are rapidly growing in popularity and the prices are coming down more than previous hardware (like modulo iPhones). For example, in a single small smartphone, there are arrays of loaded packages of technologies that are included like Wi-Fi, NFC, GPS, compass, cameras (video and still), 3-D accelerometers, Bluetooth, gyroscope, and proximity sensors and diverse information that are available through sensor fusion [10]

13. *Wizard of Oz:* The Wizard of Oz (WOz) testing is another common method of usability testing, which was developed to assist in testing low-fidelity prototypes on intended users [2]. This method allows users to cooperate with the sample prototypes in a similar manner, as they would interact with the actual interface [13]. This type of usability testing has gained recognition among agile test teams due to its

increased benefits, such as its short time-frames between releases, less trouble and cost involved in gathering participants, and the cost of execution [32].

14. ***Dairy/ Camera Studies:*** This method enables the test participants to record and illustrate certain area in their daily activities relevant to the product or software in a mechanism (diary or camera) [29]. This is a special longitudinal study and is suitable for data that are easy for test participants to record the usability problem they experienced during their interaction with software or product user interface [8, 74].

15. ***Cognitive Walkthrough:*** This is an example of expert-based usability evaluation method [28] developed to provide a team of designers the opportunity to do a quick mock-ups designs evaluation [57]. This method is considered as a task-oriented walkthrough formulated on a recognized cognitive model of novice user behaviour rather than the ease of learning analysis. This is an expert-based quantitative analysis of user's behaviour and assessing the time required for tasks' completion while interacting with the user interface of the software or application [72].

16. ***Remote Asynchronous Testing:*** Remote asynchronous or unmoderated testing [1, 47] is a situation where the moderator and the test participant are placed in a different remote location and are separated in both time and place. In remote testing, usability testing can be executed by enabling users to log onto a computer at the developer's site remotely and to carry out tasks on a site under construction [58]. In this approach, a test is conducted remotely with the assistance of prevailing communication technologies [54]. A widely used technology is the satellite video conferencing, which enables the tester to conduct test cases over the internet in a remote environment. One of the advantages of this method is that it allows the test participant to perform the usability test in a natural environment without any interference by any moderator.

17. ***Web Analytics:*** Web analytics are methods used to measure users' behaviour on the website, which is collected automatically through large samples or complete visitor populations [22]. Among the data that can be collected are the total number of the website's visitor or traffic, their demography, the number and type of links they click and the page views among others, can be analysed and reported. Web analytics are very important tools and relevant to usability and user experience professionals because it gives large-scale behaviour about the perceptions of users of the website in order to optimize the website [9]. However, this method cannot be used to know about the user motivations or the primary goals and needs. This method can only show that users of the website are leaving the checkout process but cannot give reasons behind their actions [46]. The usability issues discovered through web analytics leads to in-depth understanding necessary to find a solution to the identified usability problems. Examples include Google Analytics, CrazyEgg, Clicktale, Webtrends, Mint, and tealeaf among others [9, 65].

18. ***Remote Synchronous Testing:*** Remote synchronous testing allows the evaluator to observe the test participant in real time but they are separated geographically. Generally, remote testing involves any usability testing methods where the test participant and the evaluator are in a different location during the testing process [4, 54]. The word "remote evaluation" cut across a number of other usability methods, which collect a series of data. At times, there are little differences between remote evaluation and task-based lab testing, apart from the test participant and evaluator or moderator are in a different location [58]. Some of the main advantages of the remote synchronous testing method are that it allows the test to be done in the participant natural environment and it is easy to recruit the subjects for the test [54]. Among the tools that can be used are GoToMeeting, LiveLook, WebEx, Adobe Connect, NetMeeting, UserVue, Skype, Youguu and Glance. [4, 54].

19. ***Card Sorting:*** This method is good in information generation concerning the grouping and associations of particular data items. The test participant in a card sorting are requested to structure individual, unordered or unsorted items into different group and to label, the group base on the method used [40]. Card sorting is normally conducted at the product early design stage as a specific activity for defining an architecture but are useful during usability evaluation of the product. This approach helps to know if the identified usability issues are caused because of grouping or labelling of the groups [20]. This sorting method can be done with post-it notes or index cards or may be done automatically with other software packages [7]. For example, this method is used in user-centred design computing, it is usually used during the site architecture development, but it is useful in workflows development, toolbars, menus, and various system design elements.

The next section discusses the proposed landscape of usability evaluation methods in the context of mobile applications.

IV   RESULTS AND DISCUSSION

4.1   Usability Evaluation Methods for Mobile Applications

To validate the proposed landscape of the usability evaluation methods, the current study seeks to answer the research question in order to explain each of the evaluation methods in the context of mobile applications.

The research question seeks to discover when to use which evaluation method during the usability evaluation of mobile applications. In addition, the study seeks to uncover the type and frequency of each of the evaluation method that is commonly used in the usability evaluation of mobile applications. The research outcomes to this question give suitable data and evidence for the importance of the proposed landscape of usability methods in the context of mobile applications.

The traditional usability evaluation methods are mostly lab-based methods. They are used in the simulation of scenarios users' activities of the evaluation on mobile devices. These methods cannot adequately represent the unforeseen factors that have impacts on mobile device usage in a natural environment. Due to the highly dynamic nature of the use of mobile devices, some of the earlier usability evaluation methods would consist of different types of usability issues and might reveal diverse usability problems. In order to uncover the frequency of the evaluation method in the proposed landscape that is commonly used in the context of mobile applications, this study conducted the review of the relevant and selected 188 mobile empirical usability studies between 2005 and 2018.

Fig. 1 shows the frequency of the general usability evaluation methods used in the literature reviews of the empirical mobile usability studies. The findings show that **Survey (62%), In-lab studies (36%), Interview (34%), Observation (20%) and Log file (11%)** are the most widely used methods, which account for more than 10 percent each of the reviewed studies. In addition, **Heuristic evaluation** has 9 percent while **Think aloud and Focus group** have 6 percent each.

Furthermore, the results of the review of the mobile empirical usability evaluation studies show that remote asynchronous, web analytics, remote synchronous and card sorting are rarely used as usability evaluation methods. The reason may be due to the general apathy that remote testing on mobile devices is difficult [4, 54].
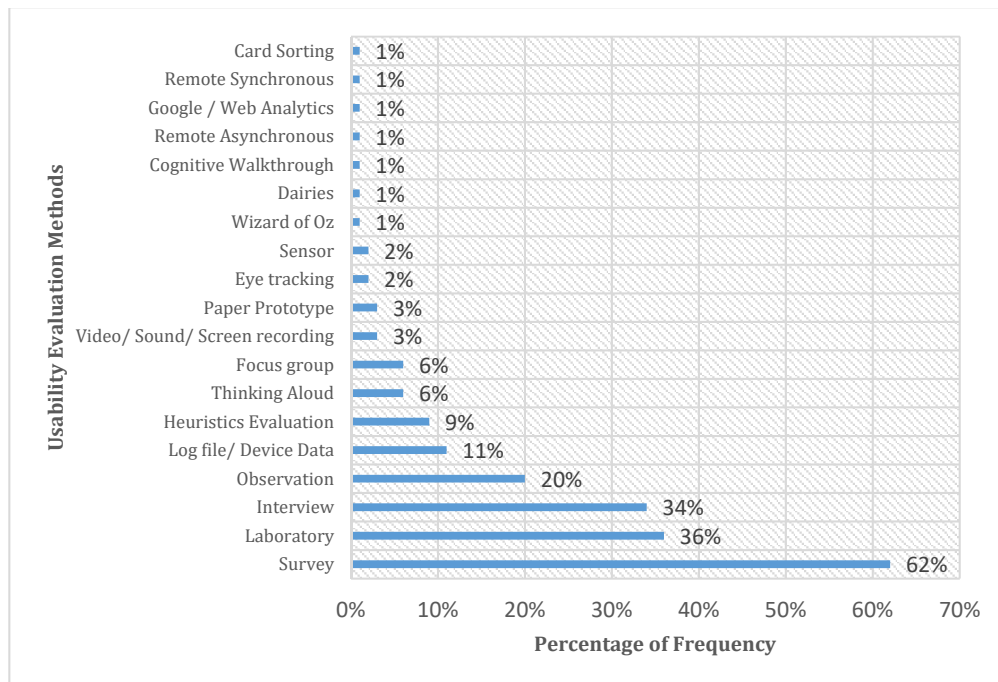

Figure. 1. Frequency of Usability Evaluation Methods Used in the Reviewed Studies

The above discussions give overviews about the 19 usability evaluation methods in the context of mobile application and as identified in the reviewed of 188 relevant and selected mobile empirical usability studies between 2005 and 2018. The next subsection discusses the proposed landscape for mobile usability evaluation methods as identified in the reviewed of the mobile empirical studies.

4.2 The Landscape of Mobile Usability Evaluation Methods

The field of user experience and usability engineering have many usability evaluation methods, ranging from in-lab usability evaluations to more recent ones like online remote asynchronous testing methods [28, 56].

It is not possible to make use of different ranges of usability methods on a specified project, it is, therefore, beneficial to combine a set of usability methods for better insight [9, 28, 56]. Surprisingly, due to the level of familiarity that usability professionals have with several these methods, they use either one or two usability methods. To have a proper understanding of what methods to use and when to use them, it is, therefore, important to categorise them using a 3-dimensional approach with axes as stated below:

- Qualitative versus Quantitative [16, 65].

- Attitudinal versus Behavioural [26, 49].

- Product context of use [62, 80].

4.2.1 The Qualitative and Quantitative Dimensions

There is a clear difference between the two dimensions, and this goes beyond the mere understanding of qualitative as open-ended questions in a research study [39]. Qualitative studies give useful and quality data about users' attitudes or behaviour based on direct observation [26, 77], while quantitative research studies gather attitude or behavioural data indirectly through the aid of certain tools like a survey or analytical tools [9]. Usability researchers have direct contact with people and observe them while they use the intended technology or product in a field or in a lab. This allows the researcher to ask the participants questions, probe their behaviour and in most cases, readjust the design or study protocol in order to meet its intended goals. The data analysis in most cases is not mathematical or statistical.
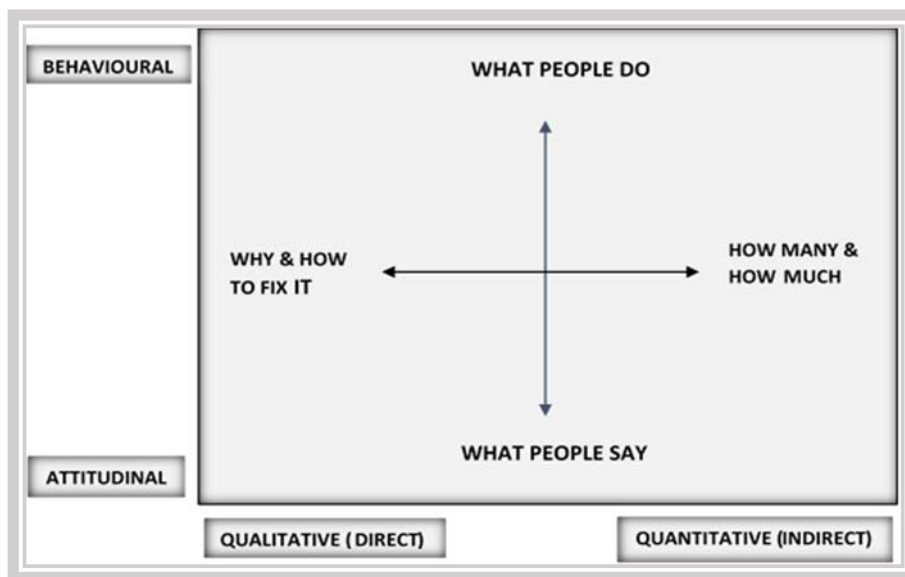

Figure 2. The landscape showing research questions answered by usability evaluation methods

However, user perceptions in quantitative research studies are generally done through mathematical analysis. The data collection instrument, like a log file or survey, gathers a large quantity of data that can be numerically coded [5, 36].

Qualitative research methods are more suitable for questions like *why* or *how to fix* usability problems, while quantitative research methods deal with questions like *how many* and *how much* in a usability evaluation [16, 65]. The huge number of data gathered assists in resource prioritization by focusing on problems that have the most impact. Fig. 2 gives general illustrations about how the two dimensions (behavioural and attitudinal, and qualitative and quantitative) affect the particular type of question.

4.2.2 The Behavioral and Attitudinal Dimensions

Fig. 3 presents and illustrates where the 19 identified usability evaluation methods appear along the dimensions stated above. The dimensions give the approach to differentiate between studies in relation to the type of questions they answer and the reasons for their suitability. The difference can be expressed by contrasting "what users do"

and "what users say" (in most cases, they are different). The goal of attitudinal research work is often to know or measure the belief of the users, hence its use in marketing departments [49].
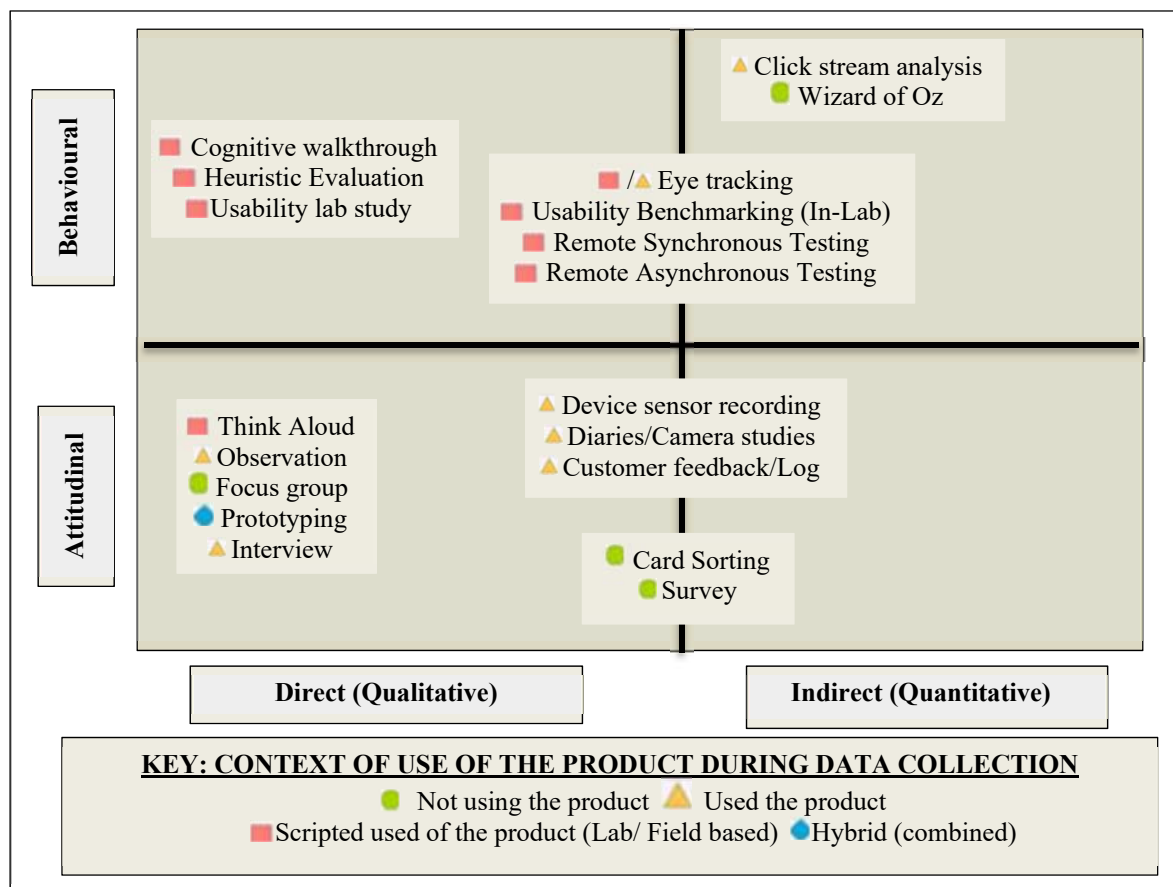

Figure 3. The Landscape of Mobile Usability Evaluation Methods

However, it is very important that usability professional focus on behavioural methods, which use self-reported data sheets of information that prove to be useful tools for designers or usability professionals [55]. For example, *card sorting* is rarely used in the context of mobile applications; it provides the mental model of a users' perception and assists in getting the best and most reliable information architecture for applications or websites [64]. *Survey* as a tool helps to measure and classify attitudes and aids in the collection of self-reported data that track major issues that require attention [41, 59]. *Focus group* is less used for the purpose of usability; however, in mobile usability evaluations, it has attracted the attention of usability professionals as it gives the view and general perception of people about the concept of the product in a group setting [10, 71]. For "what user do" with the brand or product, *eye tracking* as a tool is used to know how users interact with the product interface.

However, one of the most popular methods used is ***in-lab study*** [12], which employs a combination of behavioural and self-reported data and tends to move to either of the dimensions. In addition, it is recommended that in-lab studies should be leaning more towards the behavioural dimension.

4.2.3   Context of Use of the Product

The third dimension deals with how and if the subjects in the usability study are actually using the product or the software application. The descriptions are as follows:

- Usage of the product in ***natural*** or close to natural

- Usage of the product in a ***scripted*** form

- The study does ***not use*** the product

- The ***hybridization*** of the above

When the study involves the ***natural use*** of the product, the purpose is to know the user's behaviour and attitudes, hence interference with the study needs to be limited [3]. This approach leads to less control on the subject matter and some observational biases but gives greater validity of the findings. Examples of such methods are intercepted, analytic methods (e.g. Google Analytics, remote evaluation amongst others) and data mining [15].

Using a ***scripted*** product in a study allows the researcher to concentrate on the feelings of a particular product usage area, like a newly redesigned product. The purpose for which the study was embarked on will determine the level or degree of the scripting of the product. To get reliable usability metrics, benchmarking usability studies in nature is quantitative and needs to be highly scripted [61].

However, there are cases where the ***product is not used*** in empirical studies to observe issues that are wider than usability and usage, like wider cultural behaviours amongst others [18].

Hybrid methods employ a creative form of product usage in order to satisfy the goals of the product. For example, prototyping methods permit user interaction as well as the rearrangement of the design elements that serve as a user's product experience, which help to suggest solutions that satisfy their expectations and provide reasons for making particular selections [31, 76].

Most of the identified usability evaluation methods in Fig. 3 can move toward one or more dimensions. Some move during the study period in order to meet different goals. For example, field studies can deal with what people do (extensive observation) or what people say about a product (ethnographic interview). In addition, card sorting and desirability studies in most cases can have quantitative and qualitative versions while eye-tracking methods may be scripted and at the same time unscripted. The next section presents the conclusion of this study.

V   CONCLUSION

The fundamental purposes of website usability are to provide all necessary information and functions that users need clearly and quickly at a particular time. In contrary, a poorly designed website with poor usability will cost the owner or company with considerable operating and maintenance cost due to various negative impacts of the website. In the nowadays highly competitive information age, the poor usability on commercial websites or mobile applications will only result in serious consequences with no benefits.

Nevertheless, due to the highly dynamic nature of the use of mobile devices, some of the earlier usability evaluation methods would consist of different types of usability issues and might reveal diverse usability problems. In order to reveal the most commonly used usability evaluation methods in the context of mobile applications, this study presents the results of the review of the relevant and selected 188 mobile empirical usability studies between 2005 and 2018.

Prior research has shown that it is a good practice and beneficial to combine a set of usability methods in order to get wider insights during the evaluation processes. Surprisingly, because of the level of the familiarity of usability professionals with few methods, they use either one or two usability methods. The strong motivating question is when and what to do. In order to have a proper understanding of when to use which usability method. This study has presented the 19-identified usability evaluation in a landscape in order to have a proper understanding of when to use which usability evaluation methods. The proposed landscape is regarded as a goldmine for any usability professionals especially the newcomers to guide them in the selections and usages of usability evaluation methods in the usability evaluation of mobile applications.

REFERENCES

[1] Ahmad, W.F.W., Sulaiman, S. and Johari, F.S. 2010. Usability management system: A web-based automated system for managing usability testing systematically. International Conference on User Science and Engineering (i-USEr) (Shah Alam, 2010), 110–115.

[2] Alce, G., Hermodsson, K. and Wallergård, M. 2013. WozARd: a wizard of oz tool for mobile AR. Mobile Human Computer Interaction - Posters (2013), 600–605.

[3] Alshehri, F. and Freeman, M. 2012. Methods for usability evaluations of mobile devices. 23rd Australasian Conference on Information Systems (2012), 1–10.

[4] Andreasen, M.S., Nielsen, H.V., Schrøder, S.O. and Stage, J. 2007. What happened to remote usability testing?: an empirical study of three methods. Proceedings of conference on Human factors in computing systems (2007), 1405–1414.

[5] Bailey, R. 2014. Math Tutor: An Interactive Android-Based Numeracy Application for Primary Education. (New Zealand, 2014), 3–10.

[6] Beul-Leusmann, S., Samsel, C., Wiederhold, M., Krempels, K.-H., Jakobs, E.-M. and Ziefle, M. 2014. Usability evaluation of mobile passenger information systems. Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience. Springer. 217–228.

[7] Böhm, V. and Wolff, C. 2014. A Review of Empirical Intercultural Usability Studies. Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience. Springer. 14–24.

[8] [Bruun, A., Gull, P., Hofmeister, L. and Stage, J. 2009. Let your users do the testing: a comparison of three remote asynchronous usability testing methods. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (2009), 1619–1628.

[9] Cáliz, D. and Alamán, X. 2014. Usability Evaluation Method for Mobile Applications for the Elderly: A Methodological Proposal. Annual Conference on Human Factors in Computing Systems (Italy, 2014), 252–260.

[10] Carmien, S. and Manzanares, A.G. 2014. Elders Using Smartphones–A Set of Research Based Heuristic Guidelines for Designers. Universal Access in Human-Computer Interaction. Universal Access to Information and Knowledge. Springer. 26–37.

[11] Carta, T., Paternò, F. and Santana, V. 2011. Support for remote usability evaluation of web mobile applications. Proceedings of the 29th ACM international conference on Design of communication (2011), 129–136.

[12] Castro, L.A., Favela, J. and García-Peña, C. 2011. Naturalistic enactment to stimulate user experience for the evaluation of a mobile elderly care application. Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services (2011), 371–380.

[13] Chalkia, E., Colomer, J.B.M., de los Rios, S. and Rojo, I.C. 2014. Evaluating the Global Public Inclusive Infrastructure: Cloud4all Evaluation Framework. Proceedings of the 8th International Conference on Universal Access in Human-Computer Interaction. Design for All and Accessibility Practice (2014), 358–367.

[14] Chaparro, B.S., Phan, M.H., Siu, C. and Jardina, J.R. 2014. User Performance and Satisfaction of Tablet Physical Keyboards. Journal of Usability Studies. 9, 2 (2014), 70–80.

[15] Charfi, S., Ezzedine, H. and Kolski, C. 2014. Evaluation Based Graphical Controls: A Contribution to Mobile User Interface Early Evaluation. (2014), 114–123.

[16] Cheng, S. 2011. The research framework of eye-tracking based mobile device usability evaluation. Proceedings of the 1st international workshop on pervasive eye tracking & mobile eye-based interaction (2011), 21–26.

[17] Costa, C.J., Silva, J. and Aparício, M. 2007. Evaluating web usability using small display devices. Proceedings of the 25th annual ACM international conference on Design of communication (2007), 263–268.

[18] Coursaris, C.K., Hassanein, K., Head, M.M. and Bontis, N. 2012. The impact of distractions on the usability and intention to use mobile devices for wireless data services. Computers in Human Behavior. 28, 4 (2012), 1439–1449. DOI:https://doi.org/10.1016/j.chb.2012.03.006.

[19] Cuddihy, E. and Spyridakis, J.H. 2012. From web designer to researcher: Using experiment patterns to research web design decisions. International Professional Communication Conference (IPCC) (2012), 1–10.

[20] Dahl, Y. and Svanæs, D. 2008. A comparison of location and token-based interaction techniques for point-of-care access to medical information. Personal and Ubiquitous Computing. 12, 6 (2008), 459–478. DOI:https://doi.org/10.1007/s00779-007-0141-8.

[21] Deegan, R. 2013. Managing distractions in complex settings. Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services (2013), 147–150.

[22] Djamasbi, S., McAuliffe, D., Gomez, W., Kardzhaliyski, G., Liu, W. and Oglesby, F. 2014. Designing for success: Creating business value with mobile user experience (UX). International Conference on Human-Computer Interaction (HCI) in Business. Springer. 299–306.

[23] Façanha, A.R., Viana, W., Pequeno, M.C., de Borba Campos, M. and Sánchez, J. 2014. Touchscreen Mobile Phones Virtual Keyboarding for People with Visual Disabilities. Human-Computer Interaction. Applications and Services. Springer. 134–145.

[24] Frohlich, D., Robinson, S., Eglinton, K., Jones, M. and Vartiainen, E. 2012. Creative cameraphone use in rural developing regions. Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services (2012), 181–190.

[25] Gabrielli, S., Mirabella, V., Kimani, S. and Catarci, T. 2005. Supporting cognitive walkthrough with video data: A mobile learning evaluation study. Proceedings of the 7th international conference on Human computer interaction with mobile devices & services (2005), 77–82.

[26] Gündüz, F. and Pathan, A.-S.K. 2013. On the Key Factors of Usability in Small-sized Mobile Touch-Screen Application. International Journal of Multimedia and Ubiquitous Engineering. 8, 3 (2013), 115–138.

[27] Haapalainen, E., Kim, S., Forlizzi, J.F. and Dey, A.K. 2010. Psycho-physiological measures for assessing cognitive load. Proceedings of the 12th ACM international conference on Ubiquitous computing (2010), 301–310.

[28] Ham, D.-H. 2014. A model-based framework for classifying and diagnosing usability problems. Cognition, Technology & Work. 16, 3 (Aug. 2014), 373–388. DOI:https://doi.org/10.1007/s10111-013-0267-6.

[29] Hillman, S., Neustaedter, C., Bowes, J. and Antle, A. 2012. Soft trust and M-Commerce shopping behaviours. 14th international conference on Human-computer interaction with mobile devices and services (2012), 113–122.

[30] Holland, C., Garza, A., Kurtova, E., Cruz, J. and Komogortsev, O. 2013. Usability evaluation of eye tracking on an unmodified common tablet. CHI'13 Extended Abstracts on Human Factors in Computing Systems (2013), 295–300.

[31] Holzmann, C. and Vogler, M. 2012. Building interactive prototypes of mobile user interfaces with a digital pen. Proceedings of the 10th asia pacific conference on Computer human interaction (2012), 159–168.

[32] Hosseini-Khayat, A., Hellmann, T.D. and Maurer, F. 2010. Distributed and Automated Usability Testing of Low-Fidelity Prototypes. Agile Conference, IEEE Computer Society (2010), 59–66.

[33] Howard, G.R. 2009. Investigating good usability consistency within and across the South African super 14 rugby franchise web sites. University of South Africa.

[34] Iqbal, R., Shah, N., Chao, K.-M. and Karatsivi, K. 2008. User-Centered Design and Evaluation of WAP-Based Systems. IEEE International Conference on e-Business Engineering (2008), 528–533.

[35] Jadhav, D., Bhutkar, G. and Mehta, V. 2013. Usability evaluation of messenger applications for Android phones using cognitive walkthrough. (2013), 9–18.

[36] Jing, Y., Ahn, G.-J., Zhao, Z. and Hu, H. 2014. RiskMon: continuous and automated risk assessment of mobile applications. (2014), 99–110.

[37] Johnson, R.R., Salvo, M.J. and Zoetewey, M.W. 2007. User-Centered Technology in Participatory Culture: Two Decades "Beyond a Narrow Conception of Usability Testing." IEEE Transactions on Professional Communication. 50, 4 (Dec. 2007), 320–332. DOI:https://doi.org/10.1109/TPC.2007.908730.

[38] Jokela, T., Koivumaa, J., Pirkola, J., Salminen, P. and Kantola, N. 2006. Methods for quantitative usability requirements: a case study on the development of the user interface of a mobile phone. Personal and Ubiquitous Computing. 10, 6 (2006), 345–355. DOI:https://doi.org/10.1007/s00779-005-0050-7.

[39] Joyce, G. and Lilley, M. 2014. Towards the Development of Usability Heuristics for Native Smartphone Mobile Applications. Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience. Springer. 465–474.

[40] Keates, S. 2015. A pedagogical example of teaching Universal Access. Universal Access in the Information Society. 14, 1 (Mar. 2015), 97–110. DOI:https://doi.org/10.1007/s10209-014-0398-4.

[41] Kojo, I., Heiskala, M. and Virtanen, J.-P. 2014. Customer Journey Mapping of an Experience-Centric Service by Mobile Self-reporting: Testing the Qualiwall Tool. Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience. Springer. 261–272.

[42] Kolhe, P.R., Khetri, G.P. and Deshmukh, N.K. 2013. Study of Standard Assumptions of Graphical User Interface (GUI) Based on Usability, Adaptability and Security Factors. International Journal of Advanced Research in Computer Science and Software Engineering. 1, 1 (2013), 90–95.

[43] Koutsabasis, P., Spyrou, T. and Darzentas, J. 2007. Evaluating usability evaluation methods: criteria, method and a case study. Human-Computer Interaction. Interaction Design and Usability. Springer. 569–578.

[44] Kumin, L., Lazar, J., Feng, J., Wentz, B. and Ekedebe, N. 2012. A usability evaluation of workplace-related tasks on a multi-touch tablet computer by adults with Down syndrome. Journal of Usability Studies. 7, 4 (2012), 118–142.

[45] Lawson, S., Jamison-Powell, S., Garbett, A., Linehan, C., Kucharczyk, E., Verbaan, S., Rowland, D.A. and Morgan, K. 2013. Validating a mobile phone application for the everyday, unobtrusive, objective measurement of sleep. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (2013), 2497–2506.

[46] Lettner, F. and Holzmann, C. 2012. Automated and unsupervised user interaction logging as basis for usability evaluation of mobile applications. Proceedings of the 10th International Conference on Advances in Mobile Computing & Multimedia (2012), 118–127.

[47] Liang, H., Song, H., Fu, Y., Cai, X. and Zhang, Z. 2011. A remote usability testing platform for mobile phones. International Conference on Computer Science and Automation Engineering (CSAE) (2011), 312–316.

[48] Lim, J.J. and Feria, C. 2012. Visual search on a mobile device while walking. Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services (2012), 295–304.

[49] Liu, Y. and Li, H. 2011. Exploring the impact of use context on mobile hedonic services adoption: An empirical study on mobile gaming in China. Computers in Human Behavior. 27, 2 (2011), 890–898. DOI:https://doi.org/10.1016/j.chb.2010.11.014.

[50] López-Gil, J.-M., Urretavizcaya, M., Losada, B. and Fernández-Castro, I. 2014. Integrating Field Studies in Agile Development to Evaluate Usability on Context Dependant Mobile Applications. Proceedings of the International Conference on Human Computer Interaction (Spain, 2014), 1–8.

[51] Losada, B., Urretavizcaya, M., López-Gil, J.-M. and Fernández-Castro, I. 2012. Combining InterMod agile methodology with usability engineering in a mobile application development. Proceedings of the 13th International Conference on Interacción Persona-Ordenador (2012), 1–8.

[52] Mayas, C., Hörold, S., Rosenmöller, C. and Krömker, H. 2014. Evaluating Methods and Equipment for Usability Field Tests in Public Transport. International Conference on Human-Computer Interaction (2014), 545–553.

[53] Mazlan, M.F.B., Sivaji, A., Tzuaan, S.S. and Lokman, A.M. 2012. Enhancing the heuristic evaluation (HE) by development and validation of a collaborative design measurement system (CDMS): Collaborative design and measurement system for designers and testers. Colloquium on Humanities, Science and Engineering (CHUSER) (2012), 473–478.

[54] Miao, M., Pham, H.A., Friebe, J. and Weber, G. 2014. Contrasting usability evaluation methods with blind users. Universal Access in the Information Society. (Aug. 2014), 1–14. DOI:https://doi.org/10.1007/s10209-014-0378-8.

[55] Micallef, N., Just, M., Baillie, L. and Kayacik, G. 2013. Stop questioning me!: towards optimizing user involvement during data collection on mobile devices. Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services (2013), 588–593.

[56] Miesler, L., Gehring, B., Hannich, F. and Wüthrich, A. 2014. User Experience of Video-on-Demand Applications for smart TVs: A Case Study. Design, User Experience, and Usability. User Experience Design Practice. Springer. 412–422.

[57] Min, Q., Li, S. and Zhong, Q. 2009. An Empirical Study of M-Commerce Adoption from Usability Perspective. Eighth International Conference on Mobile Business (2009), 215–220.

[58] Mockler, J. 2014. UX suite: a touch sensor evaluation platform. (2014), 631–636.

[59] Moritz, F. and Meinel, C. 2010. Mobile Web Usability Evaluation - Combining the Modified Think Aloud Method with the Testing of Emotional, Cognitive and Conative Aspects of the Usage of a Web Application. 9th International Conference on Computer and Information Science (2010), 367–372.

[60] Neto, O.M. and Pimentel, M. da G. 2013. Heuristics for the assessment of interfaces of mobile devices. 19th Brazilian symposium on Multimedia and the web (2013), 93–96.

[61] Nuovo, A.D., Broz, F., Belpaeme, T., Cangelosi, A., Cavallo, F., Esposito, R. and Dario, P. 2014. A web based Multi-Modal Interface for elderly users of the Robot-Era multi-robot services. IEEE International Conference on Systems, Man and Cybernetics (SMC) (2014), 2186–2191.

[62] Olsina, L., Santos, L. and Lew, P. 2014. Evaluating Mobileapp Usability: A Holistic Quality Approach. International Conference on Web Engineering (2014), 111–129.

[63] Orlandini, G., Castadelli, G.A. and Braccialli, L.M.P. 2014. Ergonomics and Usability in Sound Dimension: Evaluation of a Haptic and Acoustic Interface Application for Mobile Devices. International Conference of Design, User Experience, and Usability (2014), 193–202.

[64] Otaiza, R., Rusu, C. and Roncagliolo, S. 2010. Evaluating the Usability of Transactional Web Sites. Third International Conferences on Advances in Computer-Human Interactions, IEEE Computer Society (2010), 32–37.

[65] Porat, T., Schclar, A. and Shapira, B. 2013. MATE: a mobile analysis tool for usability experts. CHI'13 Extended Abstracts on Human Factors in Computing Systems (2013), 265–270.

[66] Randolph, J.J. 2009. A guide to writing the dissertation literature review. Practical Assessment, Research and Evaluation. 14, 13 (2009), 1–13.

[67] Rivero, L., Kawakami, G. and Conte, T.U. 2014. Using a Controlled Experiment to Evaluate Usability Inspection Technologies for Improving the Quality of Mobile Web Applications Earlier in their Design. (2014), 161–170.

[68] Scholtz, J. 2004. Usability evaluation. National Institute of Standards and Technology. 8, 1 (2004), 1–8.

[69] Scholtz, J. and Consolvo, S. 2008. Towards a Discipline for Evaluating Ubiquitous Computing Applications. National Institute of Standards and Technology and Intel Research. 12, 1 (2008), 1–17.

[70] Seix, C.C., Veloso, M.S. and Soler, J.J.R. 2012. Towards the validation of a method for quantitative mobile usability testing based on desktop eyetracking. Proceedings of the 13th International Conference on Interacción Persona-Ordenador (2012), 1–8.

[71] Sieger, H. and Möller, S. 2012. Gender differences in the perception of security of mobile phones. Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services companion (2012), 107–112.

[72] Sim, G., Cassidy, B. and Read, J.C. 2013. Understanding the fidelity effect when evaluating games with children. Proceedings of the 12th International Conference on Interaction Design and Children (2013), 193–200.

[73] Sivaji, A., Abdullah, A. and Downe, A.G. 2011. Usability Testing Methodology: Effectiveness of Heuristic Evaluation in E-Government Website Development. 2011 Fifth Asia Modelling Symposium (May 2011), 68–72.

[74] Spillers, F. and Asimakopoulos, S. 2014. Does Social User Experience Improve Motivation for Runners? Design, User Experience, and Usability. User Experience Design Practice. Springer. 358–369.

[75] Swierenga, S.J., Propst, D.B., Ismirle, J., Figlan, C. and Coursaris, C.K. 2014. Mobile Design Usability Guidelines for Outdoor Recreation and Tourism. International Conference on HCI in Business (2014), 371–378.

[76] Tehrani, M., Ebrahiminejad, S., Zainuddin, N.M.M. and Takavar, T. 2014. Heuristic evaluation for Virtual Museum on smartphone. 3rd International Conference on User Science and Engineering (i-USEr) (2014), 227–231.

[77] Von Zezschwitz, E., Dunphy, P. and De Luca, A. 2013. Patterns in the wild: a field study of the usability of pattern and pin-based authentication on mobile devices. (2013), 261–270.

[78] Wyche, S.P. and Murphy, L.L. 2013. Powering the cellphone revolution: findings from mobile phone charging trials in off-grid Kenya. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (2013), 1959–1968.

[79] Xu, J., Ding, X., Huang, K. and Chen, G. 2014. A Pilot Study of an Inspection Framework for Automated Usability Guideline Reviews of Mobile Health Applications. Proceedings of the Wireless Health 2014 on National Institutes of Health (2014), 1–8.

[80] Yao, L., Liu, Y., Li, W., Zhou, L., Ge, Y., Chai, J. and Sun, X. 2014. Using Physiological Measures to Evaluate User Experience of Mobile Applications. Engineering Psychology and Cognitive Ergonomics. Springer. 301–310.

[81] Zapata, B.C., Niñirola, A.H., Idri, A., Fernández-Alemán, J.L. and Toval, A. 2014. Mobile PHRs Compliance with Android and iOS Usability Guidelines. Journal of Medical Systems. 38, 8 (Aug. 2014), 1–16. DOI:https://doi.org/10.1007/s10916-014-0081-6.