https://doi.org/10.31871/WJIR.13.5.22
World Journal of Innovative Research (WJIR)
ISSN: 2454-8236, Volume-13, Issue-5, November 2022 Pages 01-09

A Comparative Analysis of Website Usability Evaluation Techniques

Nengak Iliya Sitlong, Abraham E. Evwiekpaefe, Martins E. Irhebhude

Abstract— **Most system interfaces do not meet user intuition in terms of both context and the underlying actions thereof, hence reducing user's task execution efficiency. This translate into the amount of time taken to learn, recall and complete the procedure for certain task with respect to a given system or device. It was in the light of this that this work took a look at the existing human computer interaction (HCI) usability testing techniques with the intention of empirically establishing which among cognitive walkthrough, heuristic evaluation and user group is the most efficient in identifying and correcting system's usability problems. To create the research pathway a detail review of related work was undertaken to identify the lacuna therein for the research direction. A case study of the Nigerian Defence Academy (NDA) website was used as a specimen for administering the usability test methods and the outcome of each was documented. The research methodology adopted a purposive and stratified sampling technique to reduced biased and increased reliability of the representative sample chosen for this research. The sample chosen was an extract of final year undergraduate students of Federal College of Education (FCE) Pankshin, an affiliate of University of Jos, Plateau State, Nigeria. The students have taken a course in Human Factor in System Design, which provided them with adequate training in usability test techniques, 20 out of the total 30 students with the requisite skill were chosen to undertake the test process. The data collated passed through paired sample T-test procedure of the compare-mean analysis with confidence interval percentage of 95%, using the IBM Statistical Package for the Social Sciences (SPSS) version 26. The outcome of the T-test showed that the respective means of cognitive-walkthrough, heuristic-evaluation and user-group are 5.750, 4.900 and 6.350, with respective correlation coefficients of 0.192, 0.54 and 0.624. This shows a strong relationship between each pair of the test techniques. However, user-group usability test technique with the highest mean accuracy interval of 1.450 is the best in terms of performance relative to the other two techniques. Finally, the result of the analysis as well as charts created and the participant's findings were summarized thus proffering way forward and recommendations for the improvement of system and user interaction.**

Index Terms— **Heuristics, Human Computer Interaction, Techniques, User Group, Web Usability.**

I. INTRODUCTION

Human Computer Interaction (HCI) deals with two-way communication between a user and a system. This usually involved a communication pathway that mediates between the system and the user such as the graphic user interface that is linked to the background logic that facilitates communication process [37]. Fields of concern that led to the successes in HCI comprise of computer science, cognitive science, human factors, software engineering, management science, psychology, sociology, and anthropology as shown in Figure 1.1 [25]. Initial explorations in the field of HCI centered on concepts and interest surrounding user interface. Quintessential areas of interest in HCI relates to the attributes of input and output channels on interactive gadgets, this revolves around qualities such as interface learnability for new users as well as efficiency and knowledge replicability for regular users. HCI also focus on the provision of right and instinctive array of interaction features such as command languages, menus, and graphical user interfaces (GUI).

The paradigm shift in HCI is focusing towards patterns and the intended role by the user. The value of a system is the measure of collection of functions or benefits it render to the users. However, this value becomes obvious when the benefits or functions enhance the effectiveness and efficiency of users operations [2].

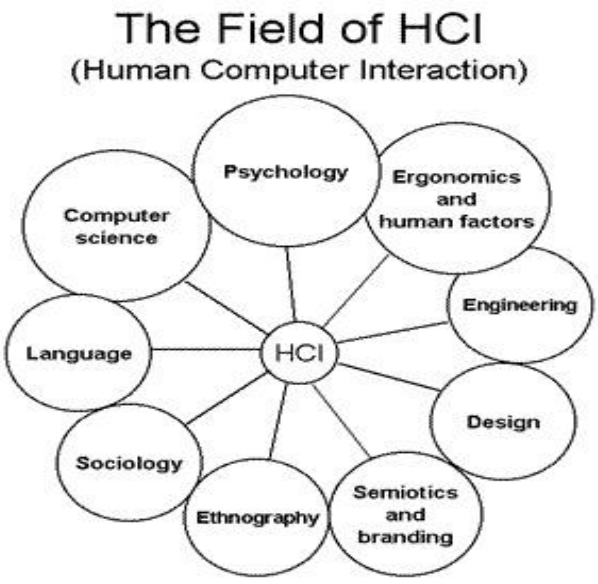

Figure 1.1: Composition of HCI Field [25].

Usability of a system that provide some services is the extend or rate at which the system can be applied proficiently and satisfactorily to achieve specific goals for a given user. The true satisfactoriness of a system is attained when there is remarkable equilibrium between the system's services and user friendliness of the system [17]. It is in line with this, that this research focused on the available usability evaluation techniques, which are user-group, cognitive walkthrough and

Nengak Iliya Sitlong, Department of Computer Science, Federal College of Education, Jos, Nigeria
Abraham E. Evwiekpaefe, Department of Computer Science, Nigerian Defence Academy, Kaduna, Nigeria
Martins E. Irhebhude, Department of Computer Science, Nigerian Defence Academy, Kaduna, Nigeria

Heuristic evaluation, to conduct a comparative study of their performance efficiency with respect to evaluating a common specimen to determine their efficacies. The outcome of this analogy will help establish which technique will be most suited for identifying and correcting user interface problems that may hinder proper human machine interactions.

The Internet has become an essential means of communication offering opportunities for carrying out myriad of activities via the World Wide Web (WWW). Particularly the internet is used in areas such as business, leisure, learning, and so forth, causing a vertiginous growth in the total number of websites in existence [12]). However, most of the websites are inaccessible to varying degrees. For this reason, several categories of users are not capable of accessing significant part of the information included in the Web. Among the class of individuals who find difficulties in accessing the Web are people with disabilities [2]. The validity period of websites is currently very short. In fact, Web designers have to manage the design of new websites or new versions of existing websites in very short time periods, which has a detrimental effect on the value and convenience of the final product [32]. Although significant research is being carried out on web accessibility, principally on the improvement of automatic accessibility evaluation tools, accessibility is not being sufficiently considered in the website development process [23]. Furthermore, during most design processes, the accessibility of a web application is usually evaluated in the latter phases of the development process, when its implementation is almost complete. As a result, correcting the detected accessibility errors implies a complete redesign of the application, which can hardly be afforded [22]. To avoid these situations developers should consider accessibility from the very beginning of the product design stage.

In addition, there is inadequacy of design methodologies that incorporate accessibility as an essential property of the product. Such methodologies should be designed and implemented within organizations in order to increase the developers' awareness of accessibility to facilitate the development of accessible websites. Not only would these methodologies lead to the production of accessible applications but it would also lead to the development of higher quality products and facilitate their maintenance [39].

A usability test of a website is commonly conducted, either to detect problems requiring improvement or to determine the usefulness of the website, whether it is serving the needed purpose as it should to the target audience [27]. Problems are likely to be discovered through the process of formative evaluation in the course of the design and implementation of a website since developers rely on usability testing to detect the problems in order to correct them [1]. A system is confirm to be usable, if the user can execute the operations required intuitively devoid of any complications, difficulties, interference, uncertainty, or confusion. This is why problems need to be identified on user interface through usability testing process and eliminate them before they are deployed for use [36]. Factors leading to website usability comprises design defects emanating from information structuring, navigation, usage of technical unpopular terminologies, design pattern of user interface and layout [10]. In other words, users should be able to perform usual operations on the website, hence, website designers ought to put in painstaking considerations during design process to be able to deliver a usable product. Usability evaluation is a useful means of ensuring that interactive systems are tailored towards user's expectation to avert errors during usage [22]. Hence, usability evaluation is an ultimate consideration in a user centered design process of any interactive system, be it a software, a web site or any information and communication technology or service.

A. Webpage Accessibility Testing with Users

A correct accessibility evaluation methodology requires testing the web application with different groups of users. This approach has the capability of identifying concrete usability obstacles for the end user. Such strategy is even more significant for achieving the overall goal of accessibility when users are involved in the development of the web application, as that will adapt the interface features with the user's view [7]. This type of test is usually carried out in controlled environments such as testing laboratories where experts can observe and collect data from users. The thinking-aloud technique, consisting of users continuously vocalizing their thoughts, feelings, and opinions while interacting with the site, is very useful since it allows the detection of barriers found by users in real time [37].

Further, various tasks can be set up in order to encourage users to browse the system, and it may be useful to collect data from this interaction so that usability parameters, such as effectiveness in completing the tasks, can be calculated. If the effectiveness rating in performing a specific task is low, its solution paths should be analyzed in order to detect any existing accessibility barrier [4]. Enquiry methods such as questionnaires and interviews are broadly used in usability testing and can also be applied to accessibility testing. The questions within these questionnaires should be designed in such a way that users' answers help evaluators to determine the most significant accessibility barriers in the system [13].

B. Quantitative Metrics for Web Accessibility

There are four properties of an accessible website that can be measured. These properties are operable, perceivable, understandable, and Robust [11]. The parameters, which have to be measured to establish the level of website accessibility according to [32], are as follows:

i. Depth of the webpage in the website where the error has occurred, this defines the proximity of the webpage from the homepage.

ii. Number of errors for each test techniques.

iii. Number of times a technique is applied

iv. The priority (or level of success criteria) for a task on the webpage.

v. The metric should generate unique value for each accessibility attribute as well as the overall accessibility value of the webpage.

vi. Type of guideline related to layout or to content on the page.

However, more other aspects needs to be put into consideration when designing a quantitative metric. In addition to the number of errors (absolute value), the number of times a guideline has been applied must also be measured [28]. This process facilitates the identification of already existing and anticipated errors. For instance, if a webpage contains only one element which is an image without any text equivalent, this webpage would be completely inaccessible. On the other hand, if one webpage has several elements and one of them is an image, which has no text equivalent, it would be more accessible than the former. This is due to the greater number of options the user has while interacting with the webpage [40].

The navigational context where the error has occurred has to be taken into account in order to measure the accessibility accurately. In some cases, this parameter is more illustrative when measuring web applications. The impact of an error in a deeper webpage within a website would be lower than the impact in a webpage which is closer to the homepage since the probability of browsing a deeper level page is lower [33]. Therefore, the deeper an error occurs, the less impact it produces. It is also necessary to distinguish between the types of guidelines. Structural guidelines are those, which refer to elements used for the layout of the webpage such as frames, tables, and embedded objects, such as applets and flash technology. Content guidelines refer to the information provided by elements such as text, images, and audio elements and should have less impact on the accessibility than the former type of guidelines [3].

II. REVIEW OF RELATED WORKS

The exponential advancement in computing has made operational HCI indispensable. HCI is a discipline that deals with the way people interact with devices or systems and the measure of how efficient systems are design for effective use with respect to user requirements and seamless manipulation [20]. The success recorded in HCI is not limited to the quality of interaction, it has spanned across different other aspects of technologies over the years. The respective technological designs are expected to specialize their interfaces to carter for individuals with different special needs as against the usual or regular interfaces [33]. Several fields of study under HCI have directed their efforts in specific researches that will eventually lead to actualization of multimodality concept, smart autonomous interfaces rather than command/action based ones, and finally active rather than passive interfaces [34]. HCI is a field of study that focuses on the design, assessing and development of interactive systems for human use and the research of occurrences around the systems [40].

HCI is interested in the development process in software and hardware that produce successful human computer interaction [39]. At the stage of device designing, the mental activities gone through by users interacting with computers should be the bottom line reason being that, generally user's qualities is not commiserate with the performance efficiency of such devices. At the same time, these devices may not possess the intelligence to adapt based on the user's reaction to the abstract scenarios [32]. However, in most cases humans emphasized the inclusion of the usual mental reasoning process in interacting with systems. Humans exhibit reliable possibilities of behaving to computers, same way as they would with other individuals [35], putting this into context, conversation between humans, the translation of logical, verbal and visible gestures, is eminent in comprehending interaction patterns. The most important essence of HCI is to enhance the communication process between users and computers. This renders computers useful and friendly to the user's desires. HCI creates or increase some areas in device implementation, some of such vital areas are: Safety, Utility, Effectiveness, Efficiency and Usability [23]. Right from 1990's, the concept of usability has penetrated the mainstream of the entire undertakings in HCI, such that it was claimed that the field of HCI is a field of Usability research [19].

A.Techniques of Usability Testing

The generally used approaches for usability testing are: cognitive walk-through evaluation, heuristic evaluation, laboratory testing, conventional user test, and field testing [41]. Cognitive walk through like the other walk through inspection approaches such as the heuristic evaluation and the activity walk through methods adopt the same approach of interaction styles. Primarily, there are two interaction styles used by the Walk-Through inspection methods, each for a certain category of users namely the recognition for inexperience users and the recall interaction styles for computer savvy users [41]. Conventional user test is used to measure the application functionality, visual ergonomics where previous knowledge is required [13]. Laboratory testing is usually conducted among users under certain environmental conditions involving staff, other participants, devices, and tasks. This method provides useful and necessary information for the evaluation process [30]. Field test provides testing guarantee of the graphic user interface workability in the actual field [9].

A satisfactory user interface layout typically requires the deployment of different usability evaluation techniques [27]. Examples of such techniques is end-user think aloud procedure and the heuristic evaluation technique. The heuristic evaluation technique is one of the most important and cost effective method among the other methods [26]. However, it requires the use of supplementary software for remote monitoring of participants. End-user think-aloud protocol technique is organized around instructing the participants to speak out audibly the thoughts running across their minds when using a website or an application software.

B. Approaches for Usability Engineering

Usability engineering is a collection of techniques to design user-friendly systems and a procedure in which developers, users, and technical designers work in harmony in order to incorporate knowledge and experience of all the stakeholders in the design process to improve the quality of the system [24]. The techniques of usability engineering is subdivided into user-focused and

expert-focused methods [11]. Within the expert-focused techniques are different versions of heuristic evaluation. Heuristic evaluation is an approach of tracing usability problems in a design, by opposing it against a set of recognized usability standards [31]. The established usability standards are listed in guidelines in Table 2.1 [28] or the "Heuristics for Web Communications." During the test procedure, experts put the system under investigation side-by-side with the guidelines and measure the compliance of the system's interface with recognized usability standards. The benefit of expert-focused process is that it is a fairly simple and fast process. A relatively small number of five evaluators can find some 75 percent of the usability problems of a product in a relatively short time [21]. The drawbacks are the rigidity that experts must be the participants in the evaluation process and they must restrict themselves within the confines of the subject knowledge, they cannot introduce new concepts outside what they already have. So they will always be substitute analyzers (expert evaluators who emulate users) [18].

Table 2.1: Heuristics of Web Usability Testing [28].

Web usability heuristics	**Description/Examples**
Visibility of current Web page status	Users need to know at each Web page as to "Where am I?" and "Where can I go next?" For example, (1) making sure each page indicates which section it belongs to. (2) internal or external links should be clearly marked; and (3) use the clear URL for distinguish the different Web pages
Match between the system world and the real world based on targeted users	Word and phrase uses on the Web site must be acquainted to the user For example, (1) a multiple-language support for the Web site of a globalized firm and (2) a Web site for children should use child-friendly phrases.
Support user control to Web navigation and relevant links	Users frequently select system functions in error and will require an obviously handy emergency recovery path to get back to the right track. For example, a "home" button on all the pages put users in total control of the website.
Consistent Web design and conformation to standards	Users should not be confused due to the ambiguity of words, situations, or actions. For example, consistent wording with links, page titles, and page headers.
Error prevention with informative contents	Even better than good error messages is a careful design which prevents a problem from occurring in the first instance. For example, (1) use JavaScript to prevent some missing parts before users submit and (2) the essential items to be filled in should be clearly indicated
Recognition rather than recall	Make icons, functions, and options conspicuous. The user do not have to memorize details from one page to be able to use another page effectively. For example, (1) if users can recognize where they are by looking at the cur- rent page, without having to recall their path from the home page, they are less likely to get lost. (2) The mouse overs for the links should be avoided
Flexibility and efficiency of use for frequent visitors	Shortcuts may often speed up the interaction for the frequent visitors. For example,(1) bookmarks and (2) single sign-on process (e.g., Amazon. com)
Aesthetic and minimal scrolling design	Dialogues should not have details that is unrelated or seldom needed. For example, put the more general information higher up in the contents hierarchy and let users scroll down deeper if they want the details.
Help users identify, detect, and recover from errors	Error messages should be expressed in plain language. Every error message should offer a solution (or a link to a solution) on the error page. For instance, if a user's search returns no result, avoid telling him/ her to broaden their search; leave a clue that will broaden the search.

Help and documentation	A system could be appreciated more if it could be used without any documentation. However, it is impeccable to provide help and documentation. For example, help pages

Thinking Aloud and Observation

Some advocates of usability testing emphasized that it is pertinent not to interfere with the user. They recommend a quiet test, where the user's activities are recorded and video taped while they attend to the task in a normal fashion [12]. Afterwards, the evaluator analyze the actions in detail. This approach may be accurate if the evaluator only desire to observe the problems, however, a considerable cases of usability evaluation focus on the discovery and correction of the problems. Observing the usability problems is one phase of the process, to trace the roots of the problems and rectify them is another phase in the cycle of usability testing. To succeed with this, it is imperative to be conversant with the expectations of the users with respect to the product under investigation, and why may not want one feature of the product or the other. This information is not available from the user's behaviour or a video tape [16]. The simplest approach is to engage the users in think aloud process during the test, or carefully probe into their intentions of the task in progress if their reactions appear to be weird. Specifically the act of thinking aloud influence the users to proceed differently, for example with a different performance [14], have investigated this issue. They evaluated four techniques:

a. Thinking-aloud: The users think aloud while undertaking the task

b. Record and thinking-aloud: User activities are recorded, and subsequently replayed with a supportive explanation by the user of why the actions were performed the way they were carried out.

c. Record and study: User activities are recorded, and studied thereafter by the evaluators.

d. Explain later: The users undertake the task, and then comment on the problems encountered afterwards.

Technique 1 and 2 appear to reveal similar problems, but technique 2 is much more time consuming. Techniques 3 and 4 did not reveal the correct problems. Technique 3 is also very time consuming [14]. It could be added that techniques 1 and 2 reveal soft performance problems. Hard performance problems, where fast reactions or motoric aspects are the issue, can be revealed by technique 2 or 3. The recommendation for all ordinary applications is to use only technique 1: Thinking aloud [5].

D.Standard and Heuristic Evaluation

Some developers believe that adherence to standards or various kinds of heuristic evaluation (for instance, design inspection) will ensure usability. It would be nice if this were true, but at present, these techniques cannot replace usability tests [3]. Standards (or style guides) improve learnability for users knowing other systems that follow this standard. However, domain-specific problems and many other problems cannot be covered by a standard [15]. For instance, no standard can specify what terms to use for domain specific concepts. Only a usability test can reveal whether the developer got it right. Several studies have shown that a check against standards only find about 25% of the problems users encounter, although they find a lot of standard violations that users don't notice [5]. In some cases, it was observed that usability problems caused by a standard. A good example is the use of modal dialogue boxes under MS-Windows. Many users complain that in order to enter the data needed in the dialogue box, they have to see the windows behind the dialogue box. But they cannot move windows around or bring other windows forward until they have closed the dialogue box [38].

Heuristic evaluation can be an expert's inspection of the design, or a check against guidelines. Surprisingly, heuristic evaluation finds only about half of the problems that users encounter [6]. Furthermore, about half of the problems reported with heuristic evaluation are false in the sense that real users do not notice these problems. Trying to remedy the false problems is a waste of development effort [29]. Generally, heuristic evaluation should only be used to detect the most obvious problems. If a problem seems dubious or is difficult to repair, let the usability test reveal whether it is important. Briefly, standards and heuristic evaluation may help, but do not eliminate the need for a usability test.

E. Research focus

The NDA website retrieved from the URL www.nda.edu.ng/home/index was evaluated for usability features. This was achieved using all the stated techniques; cognitive walk-through, Heuristic Evaluation and the User-group test techniques. The usability measurements and metrics obtained from the three test methods used serves as primary data for verifying which of the usability test techniques, is the best performing strategy for identification of usability problems of a system's user interface.

III. RESEARCH METHOD

The sample groups chosen for this research reflects a qualified representative with adequate wherewithal for executing the outlined task with high reliability level. The group was made of twenty (20) students, all of which are in their final year undergraduate of computer science programme Federal College of Education (FCE) Pankshin an affiliate of University of Jos, in Plateau state, Nigeria. The selected group of students have undertook a course in human factor in system design whose focus is to ensure students are properly grounded in the techniques and methods for identification and correction of systems interface design glitches. The techniques of usability evaluation of interest to this research includes the use of heuristic evaluation methods with its heuristics of web usability test earlier explained. The second usability testing technique adopted is the cognitive walkthrough, which uses a cyclic interaction approach in categorizing usability problems.

A Cyclic Interaction

Cyclic interaction categorized usability problems from the perspective of **Effect, Goal** and **Action** during cognitive walkthrough test process asserted as follows:

Effect-Goal problem answers questions such as, will the user in question be attempting to perform the right task as predefined in the context? So that any outcome of operation that does not tends toward a predefine goal, results in effect goal problem.

Goal-Action problem takes care of whether a webpage widget, controls or action to be taken to produce and effect is clearly visible and represented, if not then the situation results in goal action problem.
Action-Effect problem measures whether an action to be taken provides an affordance to the effect or suggest what effect should be expected by the user when the action is taken. If the action about to be executed does not relate to its effect then the situation results in Action-Effect problem [8].

The third and last technique used in studying the usability issue of NDA website is the Think aloud user group approach. Think aloud user group requires that users speak out what they feel when using a system and an observer records the reactions from the process for future evaluation and improvement of the system.

B. Efficiency of Inspection Methods

In ascertaining the efficiency of the three usability methods used in evaluating the usability of the NDA website the following approach was applied. The efficiency of a usability test technique increases with rise in the proportion of the usability problems found on the user interface by the respective technique with respect to the sample size of the evaluators partaking in the test process [30]. Therefore, the total number of usability problems generated by Heuristic evaluation, Cognitive walkthrough and User-group think-aloud techniques were tabulated and tested for efficiency using Paired Sample T-Test of compare mean analysis facilitated by IBM SPSS version 26.

IV. ANALYSIS OF RESULT

The results obtained from the number of user's errors recorded with each usability test technique as well as task efficiency rate, were tabulated and analyzed using paired sample T-Test of compare mean analysis with a confidence interval percentage of 95%, made possible through the most recent edition of the Statistical Package for the Social Sciences (SPSS), IBM SPSS version 26.0.

To further substantiate the outcome of the analysis, clustered column chart of number of errors recorded during task execution by each user per usability method was plotted and the result is as shown in Figure 4.1, in similar vein the generated outcome of the compare mean analysis is depicted in Tables 4.1, 4.2 and 4.3.

The results so obtained speaks volumes about which of the usability technique is the most efficient in terms of the effectiveness of error identification and correction within the shortest possible time as demonstrated using the NDA website. Table 4.1 showed that the result of the comparison between the outcome of glitches identified from the NDA website by cognitive walkthrough and heuristic evaluation produced a respective mean of 5.750 and 4.900, and a standard deviation of 2.3141 and 1.7137, this represents a significant margins in the performances and data distribution patterns of the two preceding usability techniques.
Inferred from Table 4.1, the mean effective performance of cognitive walkthrough against user group (think-aloud) usability evaluation technique showed that the mean performance of user-group usability test technique is substantially greater in accuracy by 0.6000 as reflected on Table 4.1 than that of cognitive walkthrough with an average performance rate of 5.750 under the same tasks and test conditions.
Notwithstanding, both cognitive walkthrough and user-group evaluation are still bonded by a slight level of correlation of up to 0.54 as shown in Table 4.3. This position is further substantiated by the result of participant number-6 in Figure 4.1, producing a tie in the number of website glitches detected for both techniques. However, there still exist a wide margin of performance accuracy and reliability between the two usability techniques as pointed out in the same Figure 4.1, which shows 12 users out of the remaining 19 users left, proved that more usability accuracy level is achieved by user-group evaluation technique than with cognitive walk through method

Table 4.1: Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Cognitive Walkthrough	5.750	20	2.3141	.5175
	Heuristic Evaluation	4.900	20	1.7137	.3832
Pair 2	Cognitive Walkthrough	5.750	20	2.3141	.5175
	User Group (Think Aloud)	6.350	20	2.6413	.5906
Pair 3	Heuristic Evaluation	4.900	20	1.7137	.3832
	User Group (Think Aloud)	6.350	20	2.6413	.5906

The affirmation of superiority of efficiency of cognitive walkthrough over heuristic evaluation is as shown in result from Table 4.2, this proves that the accuracy and performance rate of cognitive walkthrough over heuristic evaluation technique has a mean interval of 0.8500 and this confirms that cognitive walkthrough performs better than heuristic evaluation in testing the usability of systems interfaces.
However, result of Table 4.3 indicate a great correlation level of 0.192 between cognitive walk through and heuristic evaluation. A further typical practical example of this relationship is clearer from the bars of participant number-11

on the bar chart of Figure 4.1, which shows the same performance level for the two techniques, again in the overall statistics of Figure 4.1, it showed that 11 out of the 19 remaining participant shows a better performance rate in favor of cognitive walkthrough as against heuristic evaluation. Therefore, even though both cognitive and heurist evaluation techniques exhibit some level of relationship, cognitive walk through is still ahead of heuristic walk through in terms of performance efficiency.

Table 4.2: Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Cognitive Walkthrough - Heuristic Evaluation	.8500	2.6011	.5816	-.3674	2.0674	1.461	19	.160
Pair 2	Cognitive Walkthrough - User Group (Think Aloud)	-.6000	3.6041	.8059	-2.2868	1.0868	-.745	19	.466
Pair 3	Heuristic Evaluation - User Group (Think Aloud)	-1.4500	2.0641	.4615	-2.4160	-.4840	-3.142	19	.005

Table 4.3: Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Cognitive Walkthrough & Heuristic Evaluation	20	.192	.416
Pair 2	Cognitive Walkthrough & User Group (Think Aloud)	20	-.054	.822
Pair 3	Heuristic Evaluation & User Group (Think Aloud)	20	.624	.003

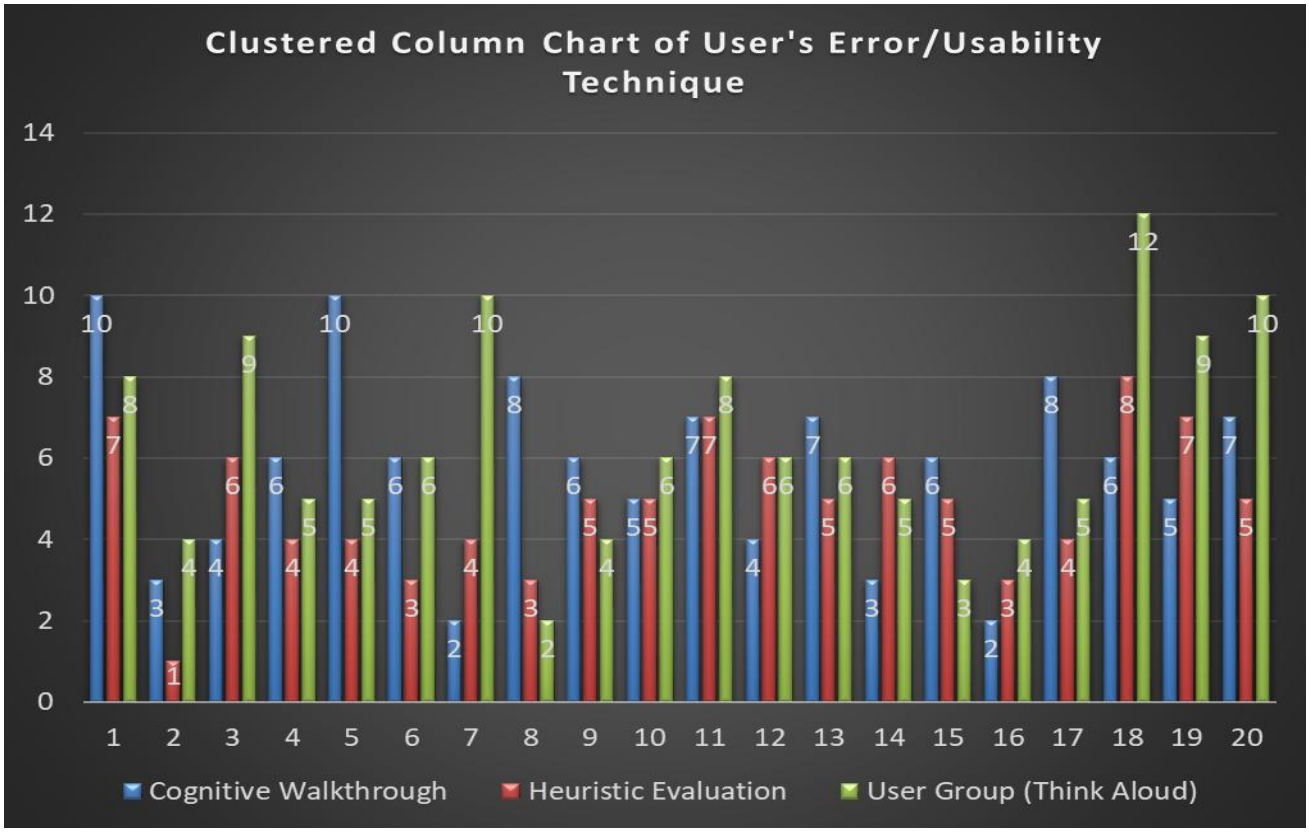

Figure 4.1: Errors Recorded by Users during Task Execution per Usability Technique

Finally, by inference since cognitive walkthrough technique is more reliable and efficient than heuristic evaluation method, and in turn user-group method supersedes cognitive walkthrough in all the usability criteria used, it then imply that user-group is the best of all the three usability test techniques employed in the evaluation of NDA website.

IV. SUMMARY

Summarily, cognitive walkthrough technique is more reliable and efficient than heuristic evaluation method, and in turn user-group method supersedes cognitive walkthrough in all the usability criteria used, it then imply that user-group is the best of all the three usability test techniques employed in the evaluation of NDA website

A.Conclusion

After painstaking planning and comparison of the major usability test procedures for system interface improvement comprising cognitive walkthrough, heuristic evaluation and user-group (think-aloud), it was established through empirical evidences that user-group usability test with a mean performance efficiency of 1.4500 is the best technique relative to its next counterparts cognitive walkthrough evaluation method with a mean performance efficiency of 0.8500 as deduced from the research conducted. Lastly, heuristic evaluation technique was found to be the least performing usability testing of the three techniques with mean performance efficiency rating of 0.6000 as obtain from the research conducted. However, the three techniques proved a significant level of performance relationship among them with the three techniques user-group, cognitive walkthrough and heuristic evaluation having respective correlation coefficients of 1.92, 0.054 and 0.624. Hence, user-group method is the most recommended and reliable technique to deploy to achieve efficiency and maximal system friendly interface that will enhance user productivity and output seamlessly.

B Recommendations

It has been established that the three usability test methods have deferent levels of performances and the one with the highest efficiency and reliability level in evaluating system's interface was found to be user-group method. Consequently, this work strongly recommends user-group technique for effective system interface evaluation and improvement whenever the need arises.

REFERENCES

[1] Aziz, N. S., Kamaludin, A., & Sulaiman, N. (2015). Assessing web site usability measurement. *IJRET*, *2*(9), 386–392.

[2] Chai, J. Y., Hong, P., & Zhou, M. X. (2004). *A probabilistic approach to reference resolution in multimodal user interfaces.* Proceedings of the 9th international conference on intelligent user interfaces. Madeira, Portugal, 5(3), 16-21.

[3] Cuomo, D. L. & Bowen, C. D. (2016). Understanding usability issues addressed by three user- system interface evaluation techniques. *Interacting with computers*, *6*(1), 86-108.

[4] Danino, N. (2014). Human-computer interaction and your site. *International journal of scientific & technology research*, *3*(6), 14-23.

[5] Desurvire, H. W., Kondziela, J. M., & Atwood, M. E. (2016): *What is gained and lost when using evaluation methods other than empirical testing*. Proceedings of HCI 16. New York, 2(2), 65-72.

[6] Diaper, D. & Lindgaard, G. (2015). West meets East: Adapting Activity Theory for HCI & CSCW applications. *Interacting with Computers*, *20*(2), 240-246.

[7] Dix, A., Roselli, T., & Sutinen, E. (2006). E-learning and human-computer interaction: Exploring design synergies for more effective learning experiences. *Educational Technology & Society*, *9*(4), 1-2. 38

[8] El-firjani, F. M. & Maatuk, A. M. (2017). A Method for Website Usability Evaluation: A Comparative Analysis. *International Journal of Web & Semantic Technology (IJWesT)* 8(3), 2-9. http://doi.org/10.5121/ijwest.2017.8301

[9] Elzakker, C. (2008). Field-Based Usability Evaluation Methodology for Mobile Geo- Applications. *Cartographic Journal*, 45(1), 139-149.

[10] Gardner, J. (2007). Remote Web Site Usability Testing - Benefits Over Traditional Methods. *International Journal of Public Information Systems*, *7*(2), 9-15.

[11] Harms, I. H. & Schweibenz, W. (2017). Usability Engineering Methods for the Web. *International Journal of Public Information Systems*, *7*(2), 21-25.

[12] Henderson, R. P., Smith, M., & Varela, A. H. (2015). An examination of four user-based software evaluation methods. *Interacting with Computers*, *7*(4), 412-432.

[13] Hertzum, M. (2015) *User testing in industry: A case study of laboratory, workshop, and field tests*: Proceedings of 5th ERCIM workshop on "user interfaces for all", New York, NY, 4(3), 6-13.

[14] Hoc, J. M. & Leplat, J. (2017). Evaluation of different modalities of verbalization in a sorting task. *International Journal of Man-Machine Studies*, *18*(2), 283-306.

[15] Jeffries, R., Miller, J. R., Wharton, C., & Uyeda, K. M. (2014). User interface evaluation in the real world: A comparison of four techniques. CHI 14 Proceedings, New York, 3(2), 53-59.

[16] Jorgensen, A. H. (2014). Thinking-aloud in user interface design: a method promoting cognitive ergonomics. *Ergonomics*, *33*(4), 501-507. 39

[17] Kantner, L. & Rosenbaum, S. (2012). *Usability Studies of WWW Sites: Heuristic Evaluation vs. Laboratory Testing*. Proceedings of the 15th International Conference on Computer Documentation. Amsterdam, NY, 6(4), 102-111.

[18] Latifi, M., Bayat, A., & Sohrabi, R. (2016). Determining the Influence of Human-Computer Interaction on Management Information Systems. *Communications on Advanced Computational Science with Applications*, 2(1), 16-23. http://doi.org10.5899/2016/cacsa-00051

[19] Levi, M. D. & Conrad, F. G. (2018). A Heuristic Evaluation of a World Wide Web Prototype. *Interacting with Computers*, *7*(4), 51-61.

[20] Lim, C. (2010). Usability testing: A review of some methodological and technical aspects of the method. *International journal of medical informatics* *79*(10), 18–23. http://doi.org/10.1016/j.ijmedinf.2008.12.004

[21] Long, M. & Dowell, B. (2016). *Measuring the Usability Index of Your Web Site*. CHI '16 Conference. Los Angeles, California, 5(2), 47-54.

[22] Long, M. & Dowell, B. (2016). *Measuring the Usability Index of Your Web Site*. CHI '16 Conference. Los Angeles, California, 5(2), 47-54.

[23] McMenemy, K. & Ferguson, S. (2007). A Hitchhiker's Guide to Virtual Reality. *Electronic Journal of e-Learning*, *2*(4), 40-49.

[24] Melorose, J., Perroy, R., & Careas, S. (2015). Usability Measurement: A Roadmap for a Consolidated Model. *International journal of scientific & technology research*, *3*(6), 35-41

[25] Moftah, R. A., Maatuk, A. M., Plasmann, P., & Aljawarneh, S. (2016). *An Overview about the Polymorphic Worms Signatures*. ACM Internatinal Conference on Engineering & MIS. New York, NY, 3(3), 53-58. http://dx.doi.org/10.1145/2832987.2833031

[26] Moftah, R. A., Maatuk, A. M., Plasmann, P., & Aljawarneh, S. (2016). *An Overview about the Polymorphic Worms Signatures*. ACM Internatinal Conference on Engineering & MIS. New York, NY, 3(3), 53-58. http://dx.doi.org/10.1145/2832987.2833031

[27] Molich, R. & Nielsen, J. (2014). Improving A Human-computer Dialogue. *Communications of the ACM, 33* (3), 338-348.

[28] Neal, A. S. & Simons, R. M. (2015). Playback: A method for evaluating the usability of software and its documentation. *IBM Systems Journal, 23*(1), 82-96.
[29] Nguyen, T. (2012). Usability Evaluation of Web-Based Learning System. *International Journal on E-Learning, 11*(3), 281-316.
[30] Nielsen, J. (2017). Usability Testing. Handbook of Human Factors and Ergonomics. *Communications of the ACM, 33* (3), 204-113.
[31] Nowell, P. & Card, M. (2017). *Expressing voice-based emotions in mobile interfaces*. The 13th International Conference on Humans and Computers. Aizu, Fukushima, 2(2), 28-31.
[32] Nowell, P. & Card, M. (2017). *Expressing voice-based emotions in mobile interfaces*. The 13th International Conference on Humans and Computers. Aizu, Fukushima, 2(2), 28-31.
[33] Peslak, A. (2005). A Framework and Implementation of User Interface and Human-Computer Interaction Instruction**.** *Journal of Information Technology Education*, 4(1), 2-3.
[34] Reeves, B. & Nass, C. I. (2013). The Media Equation: How People Treat Computers, Television and New Media Like Real People and Places. *Journal of Medical Systems, 43*(7), 1-11.
[35] Rubin, J. & Chisnell, D. (2008). How to Plan, Design, and Conduct Effective Tests. *IBM Systems Journal, 20*(3), 43-51.
[36] Sivarajah, K., Achchuthan, S., & Umanakenan, R. (2014). Pay satisfaction and Financial Stress: A study of the school Teachers in the Northern Province, Sri Lanka. *Journal of Education and Practice*, 5(7), 16-23.
[37] Te'eni, D. (2006*). Designs that fit: an overview of fit conceptualization in HCI*, (3rd Ed*).* Human-Computer Interaction and Management Information Systems conference. Boston, Massachusetts, 6(2), 94-112.
[38] Tufte, D. (2014). *E-learning indicators approach to developing e-learning software solutions.* The International Conference on Computer as a Tool. Oakland, California, 2(2), 7-15.
[39] Vinicio, L., Díaz, M., Arturo, J., & Perezgil, R. (2010). *Structured User Interface Design as a Lateral Thinking Tool for User Interface Design***.** IEEE Computer Society. Villa Nueva, Guatemala, 6(3), 70-85.
[40] Zaid, B., Jamaludin, R., & Wafaa, B. (2012). A Comparative Study of Usability Methods for Mobile Applications. *International Journal of Scientific & Engineering Research* 3(8), 2- 4.
[41] Zaphiris, P. & Kurniawan, S. (2006). Human Computer Interaction Research in Web Design and Evaluation. *Journal of Education and Practice*, *2*(2), 24-32.