Remote Asynchronous Testing: A Cost-Effective Alternative for Website Usability Evaluation

Rob Martin[1], Majed Al Shamari[2], Mohamed E.Seliaman[3] and Pam Mayhew[4]
[1,4] School of Computing Sciencesn
University of East Anglia
UK
[2,3] Department of Information Systems
College of Computer Sciences and Information Technology
King Faisal University, KSA
Email: smajed {at} kfu.edu.sa

Abstract—**The current literature suggests that the performance of in-lab and remote usability testing methods results in similar findings. However, developers still view usability testing as an expensive process that requires streamlining and improvement. This paper attempts to discover if remote testing is a viable alternative to laboratory based testing and to evaluate the costs of laboratory and remote usability testing to confirm the cost benefits of remote asynchronous testing. Therefore, a remote asynchronous usability evaluation was conducted on a selected website while a laboratory-based usability evaluation was also conducted as a control group. The results reported that , remote testing can be considered of comparable or higher quality than laboratory testing when analyzing the similarities and differences of both testing methods. Remote testing was also reported to use up to 15 days less direct evaluator/consultant involvement, which resulted in cheaper costs per problem.**

Keywords- remote usability testing, in-lab testing, usability testing cost

I. INTRODUCTION

Usability is an essential consideration for software development methods in the modern age of application and web utilization. Several methods have been developed for usability professionals to conduct evaluations [1]. Despite this IT companies are still unwilling to dedicate enough resources to usability evaluations [2]. Traditional laboratory based usability evaluation has become a reliable but expensive testing procedure. This approach for usability testing has been well documented to find 80-85% of usability problems by testing as few as 3-5 participants [3]. However, it suffers from the disadvantages of being costly and time consuming [4]. The relatively undeveloped area of remote usability testing has the potential to be more cost-effective while producing comparatively high quality results over a much larger sample size [5]. This has lead businesses and academics to consider remote usability testing as an "increasingly important alternative to conventional usability testing" [6] and [7]. Remote methods allow evaluators to concentrate on processing results, reduce laboratory costs and allow testing in the user environment. "Success of using remote methods is due to two important factors: they reach real users remotely located and consider the WWW idiosyncrasies" [8]. Remote usability evaluation is divided into two distinctive types; synchronous and asynchronous methods. Synchronous testing separates the user from the evaluator spatially but not temporally. It could potentially be done over much larger distances and therefore allow evaluators to use laboratory testing methods without the necessity of the user being in the same location. Synchronous methods produce similar results to laboratory testing [9] but continue to keep costs high. Additional problems of trust, malfunctions, support and controlled environments also need to be considered [6].

This paper aims to further investigate the performance of remote asynchronous usability evaluation compared to its more traditional laboratory based equivalent. It pays particular attention to the relative cost effectiveness of both methods. This remainder of this paper is organized as follows. Section 3 reviews selected related work. Section 3 describes the approach used to conduct the research experiments. Results and analysis are presented in section 4. General recommendations are outlined in section 5. Finally, section 6 presents concluding remarks and highlights future research directions.

II. RELATED WORK

Asynchronous testing separates the user from the evaluator both spatially and temporally. Participants would use their own computer which assesses the user actions through testing via interactive programs or task lead questionnaires. This method takes full advantage of the cost-effective and time optimization that remote usability evaluation has to offer. However, [4] argues that the asynchronous approach does not allow for

observational data and recordings of spontaneous verbalizations to be recorded. This could potentially produce lower quality results and a lower chance of discovering usability errors. Conversely, the low costs and greater user accessibility could potentially mean many more participants could be assessed, which would produce far more conclusive and realistic results. Several studies have attempted to compare remote usability testing with the traditional usability evaluation methods in terms of their quality and costs. Reference [6] conducted an empirical study of three methods for remote usability testing and a laboratory-based think-aloud method. The three methods include a remote synchronous and two remote asynchronous conditions, where test subjects are separated both spatially and temporally. Their results suggest that the asynchronous conditions are worth conducting, although they are costly. They also report that remote usability and laboratory testing found more usability problems. They needed 24 users for their research (4 groups each consisting of 6 users). All of those users were students, with good computer and Internet experience. However, the targeted system was an email system. This arose questions regarding the characteristics of users. They did not take into consideration different age groups, different web experiences or computer skills as the targeted system is an email system so users were expected to have previous knowledge and practice of emailing systems. Reference [10] conducted a comparative study of synchronous and local (next door) usability studies for an expert user interface. 20 users were recruited for each test. They reported that there was no statistical difference between these two methods in terms of quality. Their findings are in line with [9] and [11]. They suggested further research needs to be done in order to build a knowledge base of research to understand the differences of remote and local usability tests. Reference [12] compared traditional lab testing and synchronous remote testing in terms of their influence upon participants' levels and subjective testing experience. The interesting contribution of their research is that they have shifted the focus from performance comparisons to subjective assessments. Recently, [13] have presented a low-cost method of using eye-tracking to perform remote usability testing. This method requires users to perform many actions before the actual usability test starts. In addition, it requires evaluators to perform calibration many times to obtain eye-tracking data with quality precision. Still they managed to obtain the precision which is about two times worse than in professional lab eye-trackers. They concluded that this method did not help in terms of effectiveness or efficiency. Though, it can be a foundation for further development. Reference [5] conducted a study to compare the effectiveness of synchronous remote testing approach, with the traditional lab approach. The results showed that remote synchronous method is as effective as the traditional lab method. Reference [14] showed in their study that remote usability testing using virtual lab provides similar results as of the traditional lab method. However, they did not take into consideration costs differences. Overall, the literature suggests that remote and in-lab usability testing methods usually offer very similar results. Any differences between the results of the two methods could be due to research design, number of users, the targeted website and the users' characteristics as well as other contributing factors. Reference [15] reviewed and examined several experiments designed to compare usability evaluation methods. Their examination results pointed out to serious questions regarding the efficiency and quality of those experiments in comparing the different usability evaluation methods.

III. Approach Taken

The objectives of this paper are to discover if remote asynchronous testing is a viable alternative to laboratory based testing and to evaluate the costs of each approach in order to explain the cost benefits of remote asynchronous testing. In order to achieve these goals, a website portal was designed and implemented using a Java Servlet to facilitate and guide the asynchronous remote participants through the testing process. This platform provided centralized real time support without the need for a physical human presence in the form of an evaluator. The portal included pre-test instructions, FAQ, contact details and a link to the testing website and an online questionnaire. The experiment needed a laboratory-based discussion guide and remote testing questionnaire. The discussion guide/online questionnaire was based on the discussion guide supplied by one of the largest user experience design companies in Europe, which we will refer to as a Usability Company (UC). UC is a usability consultancy company based in the UK. In order to maintain a reliable and relevant usability evaluation, collaboration with UC is undertaken. They conduct high quality laboratory based usability testing using advanced techniques such as eye tracking and key logging software. Their industry standard results could then be compared to the results of our own remote testing to analyze comparative quality. They were able to provide a 'discussion guide' which they used for testing. The discussion guide is designed to prompt the evaluator in leading the participant through the exercise. The testing sample needs to be large enough to discover the usability problems while adhering to the project constraints. Reference [16] argues that 5-6 users will be sufficient to find 85% of the problems, which includes critical problems. However, more recent research suggests that the more participants used the more conclusive the results. Therefore, 6 users were recruited for the laboratory testing, 3 of them are expert (users who understand software development) and the other 3 are casual users (the average computer users). This covers the 5-plus recommendations by [16] and [3] to find 80-85% of usability problems. Remote testing has the advantage of larger sample sizes at relatively similar costs. Therefore, 20 participants were used during remote testing, consisting of 10 expert and 10 casual users. This will satisfy the requirements reported by both [16] and [3] while emphasizing the advantages of remote testing, such as larger sample groups at no extra costs. During the evaluation, all participants performed the same tasks because the task assignment is reported to have a significant effect on usability testing [17].

Hence, the laboratory based testing was carried out in the traditional way in the university setting, with the user set a series of tasks whilst being encouraged to think-aloud and being observed by the researcher. On the other hand, the

asynchronous remote users were at liberty to perform their testing, with the aid of the website portal and online resources, wherever and whenever they desired.

IV. Results Analysis

Analysis of the usability results aims to conclude on the most cost-effective method of usability testing between remote asynchronous testing and the more traditional laboratory-based testing. However, in order to assess the cost-effective potential of remote testing, two key areas of assessment must be considered. Firstly, we must examine whether or not remote usability testing is a viable alternative to laboratory testing. This was achieved by assessing the quality of test results. Secondly, confirmation of the cost-effective benefits of remote asynchronous usability testing can be confirmed (or refuted), depending on whether comparative or higher quality results were produced. This can be achieved by evaluating the financial, time and effort expenses to prepare, conduct and evaluate both of laboratory and remote usability testing techniques.

A. Comparison of Testing Results

The problems found and their severities were highlighted as quality metrics. For example, if a problem was found that distracted the participant but did not obstruct them in the successful completion of a given task, this was classified as a minor problem by the evaluator. Should there be a problem that would severely disrupt or prevent the participant from completing a given task, this would be categorised as a critical problem.

A total number of 17 usability problems were discovered by the two testing methods. 10 of them were critical and 7 were classified as minor problems. It can be seen in Table 1 below, that remote asynchronous testing was able to reveal approximately 59% of the total number of usability problems found. Remote asynchronous was also able to find 60% of the total revealed critical problems. The Fisher Exact Test was carried out and proved that there is no significance difference between group performances in terms of identifying critical problems. The p value reported by the test was 0.328. The same test was also used to compare the performance of the two testing methods in revealing minor problems. Again, the p value calculated as 0.5 reported that there is no statistical difference between the two groups in terms of revealing minor problems. These statistical results, which show there are no statistical differences between the two methods performance, are supporting findings from previous research. For example, [6] showed that synchronous remote testing out-performed in-lab testing in identifying unique major and catastrophic problems. However, their findings were not supported by enough statistical evidence. Also, analysis by [11] did not provide any statistically significant difference between lab testing and remote testing in revealing usability problems. Reference [10] concluded that in-lab testing found more usability problems than synchronous remote testing but they did not report any statistical differences between the two methods. From this reviewed literature we can also see that the performance of in-lab and remote usability testing methods results in similar findings.

TABLE I. Lab. and Remote Asynchronous Testing Results

Testing	**Critical**	**Minor**	**Total**	**No. of unique problems**
Laboratory	4 (40%)	3 (43%)	7 (41%)	3
Remote	6 (60%)	4 (57%)	10 (59%)	6
Both	10 (100%)	7 (100%)	17 (100%)	-

The overall results from our laboratory and remote usability testing highlights the critical, minor and matching problems discovered. Similarities during testing highlight the comparative quality of both testing methods. Among the (17) problems found using both laboratory and remote testing methods, four were matching problems. 3 of these matching problems were critical and 1 was minor. The single matching minor problem was realized when using different computer screens in the secure university computer laboratories and a remote participants' laptop. Smaller screens cut off the information at the bottom of the page obstructing an alternative route option. This prevented participants from viewing the advertised links and resorting to tabs.

All of the three matching critical problems were centered around route navigation. The evaluator noticed that this had a resulting effect which created two more critical problems, such as confusing page links and vague/over-detailed search results, when trying to find an alternative route. As these key critical problems were discovered by participants during both testing methods, remote asynchronous testing results can be considered of comparable quality to laboratory testing results when analyzing the similarities between these usability methods. Unique problems found during testing underlined the differing qualities of laboratory and remote usability testing. There were many more unique problems discovered during testing than matching problems. However, only one unique critical problem was found while the rest found were unique minor problems. The unique critical problem found in remote testing and not laboratory testing was concerned with tab naming, similar to the matching tab naming critical problem found in lab testing. Remote testing discovered four unique minor problems while laboratory testing discovered two more. Regarding these results, remote testing can be considered of comparable or higher quality than laboratory testing, as remote participants found one more critical problem and two more minor problems than laboratory testing. The similarities and differences between both methods of usability testing favoured remote testing over laboratory testing. Four key problems were discovered by laboratory and remote testing, where the minor problem led to the discovery of a remote unique critical problem. Remote testing also discovered two more unique minor problems than laboratory testing. Judging by these results, remote testing can be considered of comparable or higher quality than laboratory testing when analyzing the similarities and differences of both testing methods.

B. Comparative Costs

The comparative costs of usability testing will confirm or deny the cost-effective benefits of remote asynchronous testing compared to laboratory-based testing. In the two following subsections, we examine and compare the cost effectiveness of both methods with respect to time, effort and financial costs.

Time and Effort

In a competitive business environment the productivity of a usability evaluator/consultant is an essential factor in the productivity of a company. The less time and effort spent designing, implementing, piloting, conducting and evaluating usability tests, the more cost-effective the testing can become. Therefore, this influences the analysis of laboratory and remote usability testing techniques. The simpler and less direct involvement the evaluator/consultant is required to commit for certain level of results, the more productive and cost-effective the usability testing can be considered. The most efficient method of usability testing can be concluded by comparing the different usability stages and the levels of direct involvement required during each stage by the evaluator/consultant. The time-scale used to represent the amount of time spent on each stage of the usability evaluations was compared to that provided by UC for usability consultants' working hours. This considers a working day (Monday-Friday) to be 7.5 hours with five working days per week. For example, piloting for both testing methods took a week to complete which translates to five working days. This was done to produce results that were directly comparable with professional usability evaluations being conducted in the commercial sector. Table 2 shows that laboratory and remote testing methods spent 15 days of developer contact time during the preparation stage, although the re-usability of the tools used must be considered. The remote testing website portal could remain largely unchanged as only the testing website links require changing per usability evaluation, providing the online questionnaire format remains the same. The discussion guide/online questionnaire could be reused as a template for conducting future testing, reducing the repeat preparation time required considerably. Re-usability would severely reduce the preparation development time required, with the exact figure requiring further research. Considering this, time and effort expenses for laboratory and remote usability testing techniques are relatively similar, whilst their re-usability aspects favour remote testing.

The testing process is concerned with conducting the usability testing, including the sourcing and instruction of usability participants and data collection. The major difference between laboratory and remote usability testing is the evaluator contact time and testing environment. Laboratory testing is more evaluator intensive as the participant must be located, guided and recorded with an evaluator present. Remote testing requires an evaluator to source the participants, but the website portal and online questionnaire guide the participants and record the data without an evaluator being present. Both laboratory and remote testing lasted for 10 days but the evaluator contact time was non-existent for remote testing. Sourcing participants for remote testing was conducted through e-mail contact and therefore can be considered mostly automated. Scripting an appropriate e-mail for eligible remote participants with an invitation and link to the website portal required less than an hour and is therefore considered negligible contact time. Direct evaluator involvement was not required until the results had been automatically collected. This was a much more efficient use of evaluator time and creates the potential for a single usability consultant/evaluator to easily conduct multiple simultaneous testing easily. Therefore, although laboratory testing and remote testing lasted for the same amount of time, remote testing was a much more efficient method when considering the actual consultants'/evaluators' needed time. Evaluating the data collected is arguably the most crucial stage of usability evaluations, which reveals the usability problems found before corrections, amendments and redevelopments can take place. Data evaluation includes the collating of usability results into tables and graphs while comparing the samples against each other. Laboratory results evaluation lasted for 10 days while remote results evaluation lasted for 5 days as the automated services of the web-based tools had already done much of the collection tasks. Laboratory results were recorded by hand by an evaluator and collated into tables and graphs using a spreadsheet program. Remote results were automatically collected via the online survey questionnaire forms hosted by SurveyMonkey(2010). Summary pages and filters quickly sorted the information into relevant data representations, such as problems found by expert participants only, causing the remote results evaluation to take half the time of the laboratory results evaluation. Usability evaluations require intensive evaluator time to comprehensively conclude from the testing results. The automated online survey questionnaire forms save time and effort while preserving and enhancing usability evaluators/ consultants' productivity. Consequently, remote testing was more efficient and productive when analyzing the usability testing results than laboratory usability testing. The time and effort required by evaluators/consultants for usability testing favours remote testing over laboratory testing according to the presented results. This is clearly shown in Table 3.

Table 3 compares the research results reported in this paper with other results reported by three published research papers. It can be seen that this remote asynchronous testing needed less time and effort than the others methods apart from [18] and [19]. The former did not include preparation, piloting and testing time. They reported only the analysis time which needed 26 days. The latter needed 27 days as this period were used for design and application of usability testing. It did not include the data analysis time and effort. This clearly shows that remote asynchronous testing is a more cost-effective method than laboratory testing.

TABLE II. TIME AND EFFORT COMPARATIVE OF LAB AND REMOTE ASYNCHRONOUS TESTING

Evaluation Method	Laboratory Testing		Remote Testing			
Task	Time (Days)	Cumulative Time Spent (Days)	Time (Days)	Cumulative Time Spent (Days)	Time Saved (Days)	Time Saved (%)
Discussion Guide	10	10	0	0	10	(29%)
Piloting	5	15	5	5	0	-
Testing	10	25	10	15	0	-
Evaluation	10	35	5	20	5	(14%)
Website Portal	0	35	5	25	-5	(-14%)
Online Questionnaire	0	35	5	30	-5	(-14%)
Total(all tasks)	35	35	30	30	5	(14%)

TABLE III. COMPARISON OF SOME CURRENT AVAILABLE PUBLISHED RESEARCHES

Study	**Time Spent (days)**	**Comments**
This research (*Laboratory Testing*)	35	It includes all the stages
This research (*Remote Testing*)	30	It includes all the stages
Reference [18]	26	It excludes preparation, piloting and testing
Reference [19]	27	It only includes design and application of UT
Reference [20]	43	Time only includes design, data collection and analysis

Financial Costs

Financial restraints define the modern business environment and dictate the extent to which a company can improve on its efficiency, effectiveness and productivity. Regardless of how efficient a process improvement may be, financial limitations will govern the development of any adopted changes. Laboratory testing was recorded to be a more evaluator/consultant intensive testing method. In comparison, remote testing removed much of the direct involvement by evaluators/consultants and was therefore less intensive. This resulted in fewer remote evaluator contact days required and caused a reduction in the total costs. Consultants at UC provided information on the daily rate they charge for usability consultation at £800.00 per 7.5 hour day. This figure can be used with the collected data to produce the total usability costs for laboratory and remote asynchronous evaluations. Participants used during this research were unpaid volunteers and therefore their costs are not reflected in the financial comparisons. Table 4 shows the amount of evaluator contact days times the daily cost of a usability consultant/evaluator per day. This produces the total cost of each usability evaluation if we assume that the evaluation was conducted in a business environment. Confirming the cost-effective benefits of remote asynchronous testing depends on the comparative costs of laboratory and remote testing. Comparing the costs by number of problems discovered by each method can identify the most cost-effective usability method. The total cost of laboratory testing is £ 28000, which means £4,000 per problem, the cost of each problem discovered by remote testing was £1,600 making a saving of £ 2,400(60%) compared to laboratory testing. This Judging by these results, when conducting website usability testing, remote asynchronous testing is more cost-effective method than laboratory-based testing.

TABLE IV. COST COMPARISONS OF LAB AND REMOTE ASYNC. TESTING

Testing Type	**Evaluator Days**	**Total Fee (£)**	**Cost per problem (£)**
Laboratory	35	28,000	4,000
Remote	20	16,000	1,600
Saving	15 (43%)	12,000 (43%)	2,400 (60%)

V. RECOMMENDATIONS

Analyzing the cost-effective benefits of laboratory and remote usability testing enables recommendations to be made for conducting future website usability testing. These suggestions are based on the data collected and the realization of which methods are the most cost-effective. This includes recommendations for sample compositions, participant experience, specific problems and development stages.

A. Sample Composition and Experiences

Remote testing should be used when comparing sample groups focused around a specific attribute, whereas laboratory testing should be used when focusing around multiple attributes. Data analysis concluded that remote testing was more cost-effective but harder to control in terms of sample

composition. If testing is based around a single attribute, such as participants' experience, remote testing should be used as the most cost-effective method. However, if multiple attributes are required for multiple comparisons, such as age, gender and cultural backgrounds, laboratory testing should be used as it allows more direct control of the sample composition. Expert participants reacted more favorably towards remote testing than laboratory testing in terms of the amount of problems found. Overall, twice as many critical problems were found by expert participants during remote testing than laboratory testing. Casual participants reacted indifferently between remote and laboratory testing in terms of the amount of problems found. However, the same amounts of critical and minor problems were found by casual participants during both testing methods. Therefore, experienced remote participants should be used whenever possible over non-experienced participants.

VI. CONCLUSIONS

The aim of this paper was to assess the quality of laboratory and remote asynchronous usability testing techniques and conclude on the most cost-effective method for conducting website usability testing. This could be achieved by justifying the comparative or higher quality of remote results compared to laboratory results and then comparing these to time, effort and financial costs. However, analysis of the similarities and differences questioned the rate of efficiency against the effective overall results. These results indicated that remote testing could discover more problems overall within the same period of time. Therefore, remote testing results could be considered of comparative or higher quality compared to laboratory testing results. After concluding this, the results were analyzed for the time and effort spent compared to the financial costs required. Remote testing was recorded to use up to 15 days less direct evaluator/consultant involvement, which resulted in cheaper costs per problem. Therefore, remote asynchronous usability testing for website usability evaluations can be considered a more cost-effective method than laboratory-based think-aloud studies according to the analyzed results of this paper.

REFERENCES

[1] M. Hertzum, E. J. Niels, and M. Rolf Molich, "What You Get Is What You See: Revisiting the Evaluator Effect in Usability Tests," Behaviour & Information Technology just-accepted (2013): 1-32

[2] B. Bendik, G. Ghinea, and E. Brevik, "Software development methods and usability: Perspectives from a survey in the software industry in Norway," Interact. Comput. 20(3),pp.75-385,2008.

[3] N. Bevan, B. Carol, C. Gilbert, J. Nielsen, J. Spool and D. Wixon,"The "magic number 5": is it enough for web testing?, " In CHI '03 extended abstracts on Human factors in computing systems. Ft. Lauderdale, Florida, USA: ACM, 2003.

[4] C. Bastien, "Usability testing: a review of some methodological and technical aspects of the method ," International Journal of Medical Informatics, vol. 79(4),pp.18-23, 2010.

[5] M. S. Andreasen, H. V. Nielsen, S. O. Schroder, and J. Stage, "What happened to remote usability testing? : An empirical study of three methods," In CHI '07:Proceedings of the SIGCHI conference on Human factors in computing systems. (San Jose, California, USA). ACM, New York, NY, USA, pp. 1405-1414, 2007.

[6] M. Sieker, A. Nielsen, H. Villemann, S. Simon, Ormholt, and J. Stage, "What happened to remote usability testing?: an empirical study of three methods," In Proceedings of the SIGCHI conference on Human factors in computing systems. San Jose, California, USA: ACM,2007.

[7] F. L. Lopes, P. M. Arezes, and J. C. Campos. "A literature review about usability evaluation methods for e-learning platforms," Work: A Journal of Prevention, Assessment and Rehabilitation , vol. 41,pp. 1038-1044,2012.

[8] A. A. Marco, M. D. S. Winckler, F. Carla, J. and L. Valdeni, "Usability remote evaluation for WWW," In CHI '00 extended abstracts on Human factors in computing systems. The Hague, The Netherlands: ACM, 2000.

[9] H. Hartson, Rex, Jos, C. Castillo, Kelso John and C. Neale Wayne, "Remote evaluation: the network as an extension of the usability laboratory," In Proceedings of the SIGCHI conference on Human factors in computing systems: common ground. Vancouver, British Columbia, Canada: ACM, 1996.

[10] A. J. Brush, A. Morgan , and J. Davis, "A comparison of synchronous remote and local usability studies for an expert interface," CHI'04 Extended Abstracts on Human Factors in Computing Systems. ACM, 2004.

[11] T. Tullis, S. Fleischman, M. McNulty, C. Cianchette, and M. Bergel, "An empirical comparison of lab and remote usability testing of websitesm" In: Proceedings of Usability Professionals Conference, Pennsylvania, 2002.

[12] A. Chris, and D. Liu , "The effect of testing location on usability testing performance, participant stress levels, and subjective testing experience," J. Syst. Softw., vol.83(7),pp.1258-1266, 2010.

[13] P. Chynal and J. M. Szymanski, "Remote usability testing using eyetracking," In Proceedings of the 13th IFIP TC 13 international conference on Human-computer interaction - Volume Part I (INTERACT'11), Pedro Campos, Nuno Nunes, Nicholas Graham, Joaquim Jorge, and Philippe Palanque (Eds.), Vol. Part I. Springer-Verlag, Berlin, Heidelberg, 356-361, 2011.

[14] K. C. Madathil and J. S. Greenstein, "Synchronous remote usability testing: a new approach facilitated by virtual worlds," In Proceedings of the 2011 annual conference on Human factors in computing systems (CHI '11). ACM, New York, NY, USA, pp.2225-2234, 2011.

[15] W. D. Gray, and M. Salzman, "Damaged Merchandise? A Review of Experiments that Compare Usability Evaluation Methods,",HCI, vol.13(3),pp. 203—261, 1998.

[16] J. Nielsen, "Why You Only Need to Test with 5 Users," In www.useit.com, 2000.

[17] B. Anders, and J. Stage, "The effect of task assignments and instruction types on remote asynchronous usability testing," In Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems, pp. 2117-2126. ACM, 2012.

[18] R. Jeffries, J. Miller, W. Cathleen, and K. Uyeda, "User interface evaluation in the real world: a comparison of four techniques." In Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology. New Orleans, Louisiana, United States: ACM, 1991.

[19] L. Law, and T. E. Hvannberg,"Complementarity and convergence of heuristic evaluation and usability test: a case study of universal brokerage platform," In Proceedings of the second Nordic conference on Human-computer interaction. Aarhus, Denmark: ACM,2002.

[20] L. Hasan, "Usability Evaluation Framework for E-commerce Websites in Developing Countries." Loughborough Loughborough University, 2009.