International Journal of Advances in Computer Networks and Its Security– IJCNS
Volume 4 : Issue 3 [ISSN 2250 – 3757]
Publication Date : 30 September, 2014

Mobile Device Usability: Towards Achieving Usability Excellence

Mohammed I. Alowais Ahmed S. Alawaji Fang-Fang Chua

***Abstract*— the increasing trend of using smartphones and mobile applications led to more competition among smartphone manufacturers and companies to attract mobile customers. Mobile users use phones to perform all types of tasks ranging from home and work related tasks to phone calls. Due to the cost and time needed to launch a new mobile application product, a lot of companies neglect the use of comprehensive usability testing method to gather the users' feedback before introducing the new product. In fact, mobile application products should be tested by real users to discover usability problems before the application is launched into market. A novel usability testing method for mobile devices has been proposed in this work to gather more accurate feedback from the users throughout the testing phase. The proposed method combines think-aloud protocol and field study to identify usability problems while gathering the users' feedback. The results show a comprehensive analysis of the testing method for 20 users after completing 10 different tasks scenarios.**

***Keywords*— Usability testing, mobile usability, smartphones, applications, ux, ui, user experience, user interface**

Mohammed I. Alowais/ Ahmed S. Alawaji
SAP
Centria bldg, 4th Floor
Olaya Road, PO Box 19319
Riyadh 11435
Saudi Arabia

Fang-Fang Chua
Faculty of Computing and Informatics, Multimedia University
Jalan Multimedia, 63100
Cyberjaya,Selangor
Malaysia

I. Introduction

Smartphone usability testing is becoming very important in the last few years. With the increasing number of mobile applications and systems available users, it is vital to keep usability tests as appropriate as possible for the mobile device and its applications. According to Gartner [10]; worldwide mobile device sales to end users totaled 455.6 million units in the third quarter of 2013. Furthermore, there are more than 775, 000 published applications in iTunes store according to Apple[8].

The traditional approaches of usability testing in Human-Computer Interaction literature do not cover the mobile device characteristics for several reasons such as the mobile device performance and capability, connectivity and the context in which it has been used. Mainly, the aspect of "mobility" cannot be tested in laboratory Bastien [5]. Traditional computer usability testing methods are not accurate usability measurements to mobile. Also, mobile devices vary from one another in term of capability, functionality and operating system. Thus, a new usability testing methodology for mobile phones is essential to address the differences in the characteristics between computer and mobile phones. A usability testing methodology which combines think-aloud protocol with field study has been proposed and tested to gain insight into users' experience.

II. Background

Testing the usability of mobile devices is a challenging task with the fast improvement of the mobile device's hardware and software. The use of smartphones is not only limited to making calls and sending text messages but also for many other daily activities. Hence, testing the usability of these devices needs to be addressed properly. There are many approaches that have been introduced over the past few years to test the usability of device, software or website. Today, same testing methods adopted for computers are applied to mobile devices which in fact may not fit the new generation of smartphones.

Measuring the usability of a mobile device using the field study test enables usability experts to monitor users while evaluating the software or the device in the real environment. The method offers the chance to understand the user's behavior while using the device in the real context of use. These features are not available in the laboratory where the usability test procedures set a priori and results are to be recorded. However, usability testing in the laboratory is still widely used among mobile devices manufacturer. Betiol and Cybis [2] compare three approaches for usability testing of mobile device these are; computer-based mobile phone emulator inside the laboratory, using a mobile phone inside the laboratory and using a mobile phone linked to a wireless camera in the field. Apparently the use of computer-based emulator inside the laboratory identifies more usability problems than the other two methods. It is important to make sure that the user is not feeling monitored or controlled while conducting the usability test, if for example the phone is fixed on tripod to video capture the user's activity then the movement of the user is limited that will probably cause discomfort to the user. This in turn will affect the reliability of the usability test results.

Another study Anne et al. [15] compares field study and laboratory testing for mobile applications. The aim of the study is to find out if it worth conducting field study tests instead of laboratory. The result indicates that there are no significant differences between the two methods in terms of the type of usability problems found and the execution time of a particular task. Although field study tests require more effort

than in laboratory, the research indicates that it reveals more usability problems compared to laboratories Duh et al. [9]. Additionally, field study test identifies more severe issues than when conducting lab tests. Unlike field study experiments, laboratories have limited experimental space which may restrict subjects' movement and make them uncomfortable when testing.

The context of use is considered one of the most important factors in usability testing for mobile devices as mobiles are portable devices used in different places like office, train station, and home. The idea of using a real environment is to uncover usability problems that cannot be identified in different context of use, by using the field-based approach it may be possible to obtain a higher level of 'realism' J. Kjeldskov et al. [1]. For example, the voice command feature may work fine at the office but it may not do the same while being tested in the street or train station. As a result, the context has large impact on determining the usability of mobile devices and should be considered when testing the usability of device.

Conducting usability tests can be done in many ways, yet one of the most effective methods is called "think-aloud". Think-aloud is a usability method in which an experiment subject "thinks out loud" while testing the device. According to Draper [14], "think-aloud protocol consists of observing a user working with an interface while encouraging them to "think-aloud"; to say what they are thinking and wondering at each moment". There are many advantages of using think-aloud protocol for mobile phones usability testing:

1) Testers can identify usability problems by listening to their participants' thoughts
2) It requires less number of subjects which leads to faster testing process.

Think-aloud method has been used in many papers and performed well in identifying usability problems. Als et al.[1] compare two usability testing methods: think-aloud and constructive interaction. The constructive interaction is a usability test in which two test subjects collaborate to conduct tasks using a computer system Nielsen and Landauer [12]. Their research experiment was conducted with children to evaluate think-aloud and constructive interaction while interacting with a mobile phone. Consequently, the result shows that there are more usability problems identified using the constructive interaction than think-aloud. Yet, think-aloud method shows an acceptable performance. However, conducting usability tests by children may not be effective as most of the phone usage is meant for adults to carry out certain tasks related to their work or business. Moreover, the percentage of children use for smartphone is very low compared to 90% of adults in the U.S. (2010) [7] consequently; we cannot rely on children to identify critical usability issues during testing, children are not able to discover this type of usability problems as they will not understand the importance of these features.

Also, Erica et al. [13] compares several methods of think-aloud protocol; a traditional protocol, a speech-communication protocol and a coaching protocol as well as silent control. The last method does not allow any "thinking out loud" as the name implies. The other three think-aloud protocols are different from one another in the way it is had been implemented with. For example, speech-communication and coaching methods give more space for feedback between tester and participant, while the traditional protocol allow only tester to encourage participant to keep talking by saying "keep talking .. Continue" and so on. The coaching protocol is where the user being coached by admin on how to complete particular task or in case he stops and cannot carry out the task. The result of the study shows that coaching protocol performs much better while the other two give similar performance. However, coaching protocol may not be considered as reliable think-aloud protocol for usability testing because the user is being couched and guided through the use of the device and hence, it is difficult to state whether or not this device is usable given the fact that the user not able to use it without guidance.

Also, another important measurement of usability tests results is the emotional state of participants Moritz and Meinel [11] the user behavior during the usability test can have great impact on the quality of the testing results. By observing the participant's emotional state during the test, we can have clear understanding of the occurred usability problems. the user emotion response while executing particular task can be categorize into several categories which can be summarized into two main categories negative and positive emotional states. recording the user emotion or mental state while testing the device allow usability experts together with the developers to track the exact point that leads to user's negative emotion, anger or frustration. Moritz and Meinel [11] examine the combination of field study with the think-aloud method. This combination result in a method which takes the participants to their usual working environment and use the laptop to carry on a set of tasks while thinking aloud instead of performing the tasks in the laboratory. The use of several emotional/ mental states to measure the users' behavior may add value to help understand the user attitude accurately towards particular functionality.

Usability practitioners may develop a combination of existing usability testing methods to suit the mobile device usability testing requirement. Furthermore, the new method could adapt context-aware, cognitive and conative aspects, and could add other test elements such as heuristic evaluation, questionnaires, and scenario-based tasks Lee and Grice [6].

These traditional methods used together to identify more usability problems in the mobile devices. The authors believe that testing the mobile usability depends heavily on the feedback relationship and thus to be able to identify the problem and the cause of it, there must be communication between the user and the test admin.

This paper proposes a new methodology for mobile device usability testing combining; think-aloud protocol, field study and emotional states. Think-aloud protocol proves its ability to reveal many usability problems in many applications as well as the field study which is not only successful in software

industry but in many other applications. The methodology is described in details in the following section.

III. Method

A. Experimental Setup

The mobile usability testing methodology aims to identify usability problems effectively and accurately. The motivation behind proposing such methodology is to address the differences between mobile and computer characteristics. This methodology covers the three stages: identifying, classifying and analyzing the usability problems. It combines think-aloud and field study together with emotional state observation to identify usability issues. During the usability testing, users will be requested to conduct several tasks on mobiles while simulating their usual working environment. For example, users will be asked to use their phone in the office, campus, street or shopping mall. While the users are using the mobile phone in these places they will be asked to speak loudly while performing the given tasks and the result including emotional state and comments will be carefully noted by tester**.**

The fig.1 illustrate the methodology steps which are setup, execution and analysis, those are explained in more details as follow:

Setup: In this phase, the admin starts preparing the first elements of the experiment like forms and tasks. The form is used to record the following elements: tasks scripts with locations, time, emotional state, and users' comments and information. Mostly, tasks and locations are designed based on the mobile device or the application processes.

Execution: The user is asked to complete task by task. Before starting the test, the admin explains the test procedure and describe the method participant should follow throughout testing. During testing, the admin monitors the participant and takes notes of the participant's progress which includes the participant's emotion and feedback. Those activities are voice-recorded for future reference.

Analysis: Upon completion of test tasks, the results are combined for detailed analysis. In this step, tester identifies the usability problems from each participant test. Then, combines them to identify how frequent each issue occurs per task and how many usability problem reoccur among users.

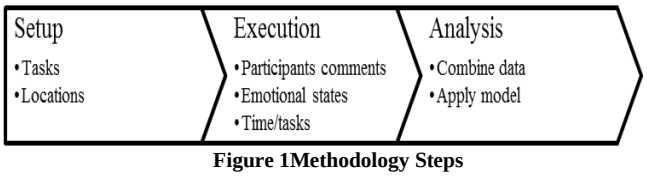

Figure 1Methodology Steps

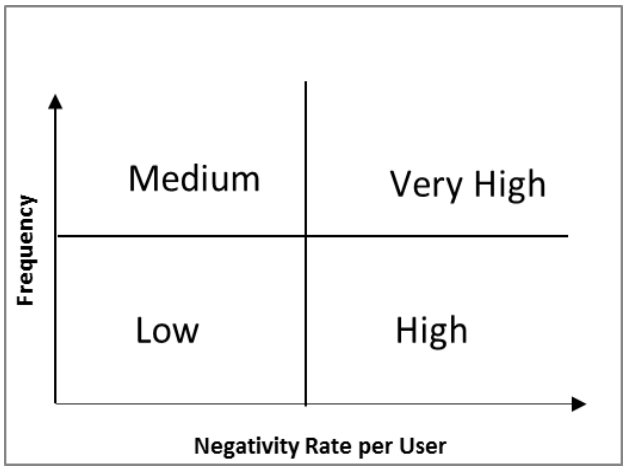

Figure 2 Severity Analysis Model

Fig. 2 shows the developed model that will be used to analyze the severity of usability issues for each task. Combining data after experiments will result in three main data elements: The total number of issues per task, the frequency of issues among users par task, and the percentage of overall negative user experience found per task. The model utilizes bubble chart to clarify these three elements for analysis. Each bubble represents a task and the size of bubble indicates the total number of issues found for a particular task. The use of this chart type is very essential to visualize and demonstrate the correlation between these three elements.

Furthermore, the model classifies each task under one of the four categories: very high, high, medium, and low. For example, if one of the tasks has resulted in many issues for participants and high negative emotions, mostly it will be located at very high section. Furthermore, bubble size will indicate the number of issues compared to the other bubbles (tasks).

B. Evaluation Steps

The focus during testing will be on the following aspects:

1. Task execution time.
2. Usability problems identified per task.
3. Usability problems frequency (reoccurrence).
4. Emotional state of the participant.

Also the participant's feedback and emotional state during the testing will be monitored for every single step in the test task to track the exact point where the usability problem has been triggered. The participants' emotional state will be assessed according to the following scale of emotional states:

1. Enthusiasm
2. Joy
3. Calmness
4. Bore
5. Anger
6. Frustration

This scale is used during the testing procedures to understand the severity and the influence of the usability problem faced by the user and to allow developers and the usability team in device manufacturer to assess the usability of their product accurately. However, to summarize those user emotional states, in this paper we split those six categories of into two types either positive or negative emotions. Defining them into two categories promote for more clarity in the results and discussion part of the research. Additionally, the participant test will be recorded for future references.

C. *System and Tasks*

A number of 20 participants aged 25 on average participated in the experiment; each of them is given a set of 10 tasks to carry them out. Within the task list/test cases there are different testing environment such as office, home and shopping mall to ensure that particular feature(s) are tested in real context of use. The exact same test set is distributed among the participants. It is also important to note that the number of tasks should be minimized to avoid complicated emotion response from the participants. The experiment carried out using Samsung Galaxy S3 which has an Android operating system. The device has several standard features like touch keyboard, voice command, handwriting, etc.

IV. Results

The main concern in this paper is to identify the issues that results in the worst user experience. From the user feedback during the execution of the test task, we were able to take detailed note of the usability problems as well as the exact stage in which the problem is triggered. Also, overall emotional states were recorded and categorized as positive or negative for each task.

The experiment was conducted on 20 participants that resulted in more than 80 usability problems. Most of the usability problems identified were frequent. Out of those 80 reoccurring usability problems 20 problems were unique. The degree of severity of usability problems was assessed into four categories; low, medium, high and very high.

Fig. 3 shows two important results: how many usability problems were found during tasks execution and how many of these problems were unique. Test tasks vary in terms of the number of usability problems identified. It is worth mentioning that the test task scenario and script will have a critical impact on the number and type of usability problems identified, the more complex the test script is the more usability problems will be identified, by having a complicated and long test script, the user will go through long steps that would increase the chances of identifying additional usability problems. However, that might affect the emotional state of participant if the overall experience was negative.

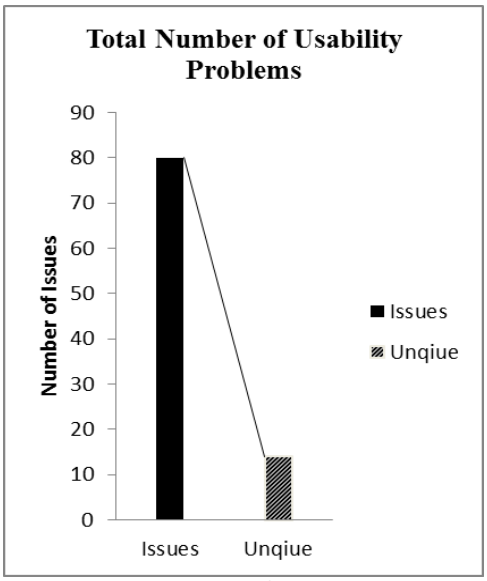

Figure 3 Total Number of Usability Problems

Also, the results indicate that the higher the number of frequent usability problem identified, the worst overall experience for users. For example, Fig. 4 compares the number of total and unique usability problem identified within tasks 3 and 4. As fig. 4 indicate that results found out about 40 issues during testing task 3 compared with only 10 usability problems identified in task 4 testing. Detailed analysis concluded that only 4 usability problems of task three were unique, whereas task four has 2 unique usability problems.

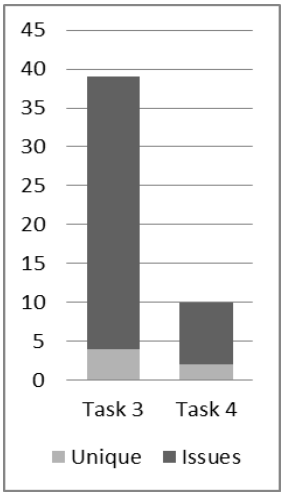

Figure 4 Usability Problem Uniqueness

From fig. 5 and 6 the result shows that 55% of users in tasks 3 identified the task experience as negative, compared with only 30% who consider the experience negative in task 4.

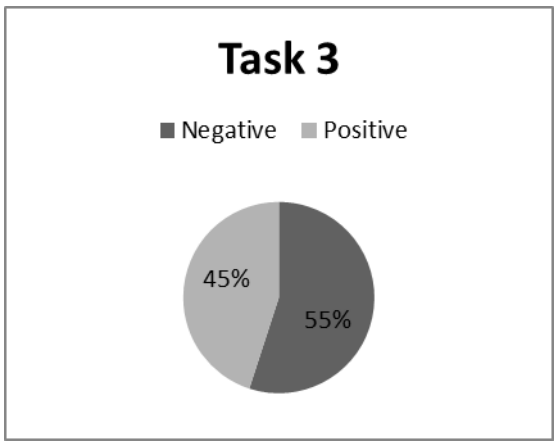

Figure 5 Task 3 Experience

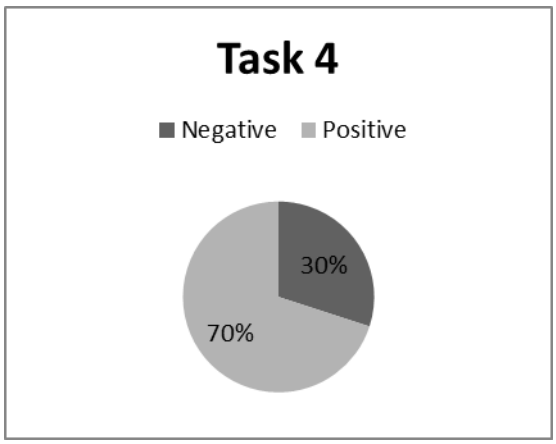

Figure 6 Task 4 Experience

In Fig. 7, the suggested model has been applied to our results. The fig. shows the result from the usability testing. There are 7 test tasks that have usability problems. The graph indicates two intervals; frequency which is calculated by finding the average usability problem per user and negativity level that is presented in percentage of task overall negativity.

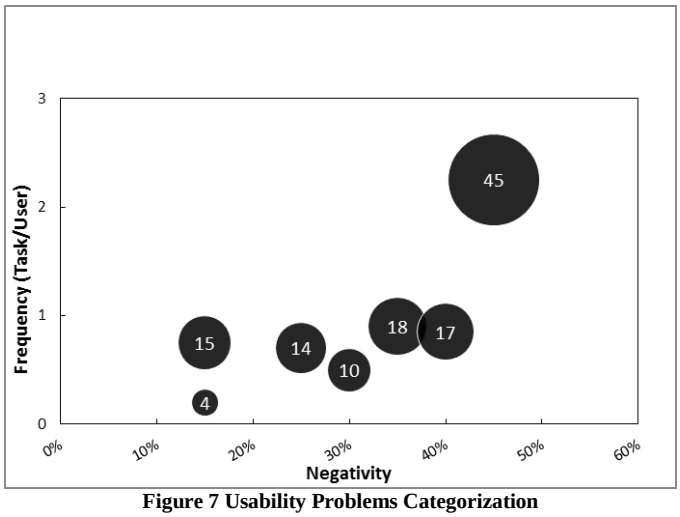

Figure 7 Usability Problems Categorization

Moving to the right side of the negativity interval means that the negativity level increases and by moving up in the frequency interval the usability problems are increasing per task per user. The size of the bubble in the graph illustrates the number of usability problems that exist for each task in the test set. The bigger the bubble the more usability problem we have for that particular task. The number inside the bubble has been shown to indicate the overall issues found in each task.

From the graph we can see that test task number 9 has the highest number of usability problems and from there it is also clear that the negativity level towards this task is very high and the frequency is high as well.

V. Recommendations

Future work can be done to evaluate the differences between usability issues results for mobile applications and systems. Also, a more detailed analysis on choosing the right locations can help when designing the usability experiment to fit the real context. Furthermore, since the scope of mental emotions is very wide, we believe that a model should be introduced to limit the number of mental states to fit usability needs. Another way to further enhance the understanding of the usability problems severity level is to experiment the use of this model with one of data mining techniques to assess the severity level of identified usability problem.

VI. Conclusion

Proper usability testing of mobile devices is very critical with the current competition among smartphone makers. Mobiles usability should be tested by users to find usability problems before releasing the mobile systems or applications. Every mobile has different functionality and capability in which it requires a unique testing method to understand users' frustration.

In this research, a methodology to test the usability of the mobile device has been introduced. The method shows good performance in identifying the usability problems in mobiles. Unlike other usability testing methods, the proposed method combines three important elements: think-aloud, emotional state and field study. The result shows that combining these three methods can fit the real mobile environment and produce a number of critical issues.

The main goal is to establish a methodology that covers all the aspects of the mobile usability testing starting from testing in real environment and ending with categorizing the usability problem according to its problem severity.

References

[1] Als, B. S., Jensen, J. J., & Skov, M. B. (2005). Comparison of think-aloud and constructive interaction in usability testing with children. In Acm Isbn (Eds.), Proceeding of the 2005 conference on Interaction design and children IDC 05(pp. 9-16). ACM Press.

[2] Betiol, A. H., & Cybis, W. D. (2005). Usability testing of mobile devices: A comparison of three approaches. (M. F. Costabile & F. Paterno, Eds.) Human Computer Interaction Interact 2005 Proceedings, 3585, 470-481. Springer-Verlag.

[3] Boren, T., and Ramey, J. Thinking aloud: Reconciling theory and practice (2000). IEEE Transactions on Professional Communication 43, 3 261-278.

[4] Ericsson, K.A. and Simon, H.A. Protocol Analysis: Verbal Reports As Data. (Revised ed.) MIT Press, Cambridge, MA, USA, 1996.

[5] J. M. Christian Bastien (2010), Usability testing: a review of some methodological and technical aspects of the method.

[6] Lee, K. B., & Grice, R. A. (2004). Developing a new usability testing method for mobile devices. International Professional Communication Conference 2004 IPCC 2004 Proceedings, 115-127. Ieee.

[7] Mobile application survey (2010), Research2Guidance, available at: http://www.research2guidance.com/

[8] Apple Press Info (2013), Apple, available at: http://www.apple.com/pr/library/2013/01/07App-Store-Tops-40-Billion-Downloads-with-Almost-Half-in-2012.html

[9] Duh, H. Been-Lirn., Tan, G. C. B. & Chen, V. H.-h. (2006). Usability evaluation for mobile device: a comparison of laboratory and field tests. Proceedings of the 8th conference on Human-computer interaction with mobile devices and services (p./pp. 181--186), Helsinki, Finland: ACM. ISBN: 1-59593-390-5

[10] SmartPhone Sales Q3 (2013), Gartner Inc., available at: http://www.gartner.com/newsroom/id/2623415

[11] Moritz, F., & 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. 2010 IEEEACIS 9th International Conference on Computer and Information Science, 0, 367-372. IEEE.

[12] Nielsen, J. and Landauer, T. K. (1993) A Mathematical Model of the Finding of Usability Problems. In Proceedings of the Human Factors and Computing Systems INTERCHI'93, ACM Press, pp. 206 – 213

[13] Olmsted-hawala, E. L., Hawala, S., Murphy, E. D., & Ashenfelter, K. T. (2010). Think-Aloud Protocols : A Comparison of Three Think-Aloud Protocols for use in Testing Data-Dissemination Web Sites for Usability. Human Factors, 2381-2390. ACM Press.

[14] Stephen W. Draper. (1998) Available at "http://www.psy.gla.ac.uk/~steve/HCI/cscln/trail1/Lecture5.html" School of Psychology University of Glasgow

[15] Kaikkonen, A., & Kallio, T. (0). Usability Testing of Mobile Applications: A Comparison between Laboratory and Field Testing.