

Mobile HCI Evaluation

Kuber Dutt Sharma
Graduate Student – CS5760
Department of Computer Science
Michigan Technological University
Houghton, Michigan
Email: kdsharma@mtu.edu

Abstract

This conceptual paper discusses some of the established industry standards and methodologies for mobile HCI evaluations. The escalated growth in use over the last decade and advancement in technology has lead mobile devices to evolve into complex multimedia interactive systems. Since mobile devices can be used in vast permutations of user settings, scientists have established the need for standard evaluation methodologies that evaluate the usability of mobile phones comprehensively and in a timely manner. Significant differences like user behavior and subjective responses to the device have been found between usability tests conducted in laboratory and real life situations.

Introduction

Mobile phones have become overwhelmingly popular in today's society. Due to the rapid development of mobile technologies and functionality the mobile phone is now regarded as a common consumer appliance (Klockar, Carr, Hedman, Johansson, & Bengtsson, 2003). The popularity of modern mobiles phones is determined by usability and user friendliness. Yet, the research of usability in mobile phones is a newly evolving area. Thus the industry is fast accepting usability engineering, for designing products. Due to the highly dynamic context of use, the conventional laboratory based usability evaluation is not a perfect fit for mobile devices. Most modern practitioners, advocate a well balanced mix of laboratory usability and field usability tests. It is important that the usability tests should test realistic scenarios, arrange the tasks by the priority, and focus on efficiency.

With extensive research, well-structured guidelines for laboratory evaluation of mobile devices have been established, it is now required that the researchers focus on developing better usability evaluation benchmarks for field evaluation of mobile devices. Another concern is that field usability testing is not always feasible. Field-based usability studies are often time consuming and the added value is questionable. This gave rise to another hybrid solution for usability evaluation, wherein scenarios requiring mobility and context simulations where added to laboratory evaluations. The two tasks at hand for modern practitioners are improving the realism of laboratory settings, and supporting high-quality data collection in field evaluations. Figure 1. depicts the rapid evolution of mobile phones.

Figure 1: Rapid evolution of mobile phones

Going Mobile | The evolution of the cellphone



HCI and Mobile Devices

According to Shackel (cited in Grudin 2009, p. 371) the core aspect of computer use has shifted “from system supremacy to personal empowerment”. Mobile devices are the driving cause of this changing trend. Researchers in HCI have moved past evaluating the single user setting. Modern HCI practitioners are devising new social and mobile application experiments. Jaimes and Sebe (2007) also emphasize that with the hardware costs rapidly going down, mobile devices will become rapidly available and complex in relation to their computational capacity.

There are four major categories of mobile devices, namely: unintelligent gadgets, cellular phones, smart phones, and devices with operating systems. With such a wide variety of devices and user situations available, mobile application developers often have to take mobile device characteristics into consideration. Some of the HCI researchers focus on adjustment of huge volumes of data on small screens, mobile usability assessment and texting. Researchers are conducting context based usability evaluations. The various contexts are: task, physical, social, temporal. Task context evaluates the performance of the actual task and also accounts for any interruptions that may occur; physical context evaluates the sensed circumstances of the situation; social context evaluates the presence of other people and their impact; temporal context evaluates factors pertaining to past or future experiences. Due to these fundamental differences between the device types, a new set of usability paradigms for mobile devices should emerge.

Related Work

The modern mobile device is highly evolved, but with the fast changing technology the life span for the devices are getting shorter. Thus it is important that the industry should standardize the development of mobile devices in order to provide the maximum usability from the device. The Kano's quality model published in 1980s, Jokela (2004) divided usability into three categories: must-have, more-is-better, and attractive usability. According to the model the "attractive usability" yesterday turns into "more-is-better usability" today and then "must-have usability" tomorrow. Most developers try to satisfy at least the "must-have usability" for user satisfaction. Some researchers suggest that if the "must-have" expectations are met, then even some serious usability problems may be accepted by the users. Usability evaluation aims at gaining a third-party assessment of user characteristics and to gauge the efficiency with which the user can access and edit content on any device (Blandford 2004). Thus it is important that a comprehensive research be done to identify the bare minimum usability features which

if provided, give users satisfaction. Various usability evaluation methods (UEMs) are being developed to assess and improve usability of interactive systems.

According to the literature on current practices of mobile evaluations, the practitioners have split opinions on the assessment of usability features using heuristic evaluation or UT. Heuristic evaluation is an established, affordable and efficient usability evaluation technique. To support the short lifespans of the mobile devices, heuristic evaluation can be performed without real end users or a working prototype in the early stage of design process. However, the lack of proximity to real life problems and user are some of the limitations of heuristic evaluation. Also, heuristic evaluation fails to provide enough information about the magnitude of the detected usability problems. Thus another hybrid approach followed in the industry is to use heuristic evaluation with cognitive walkthrough. Implementing usability analysis early in the development helps during the implementation phase. This helps in prioritizing tasks into: components required by the task; ideal components for the user; and the specified component to be implemented.

Researchers have divided UEMs into three different groups based on their functions: test, inspection and inquiry. Testing UEMs employ representative users to work on the device or a prototype and the user interface and their performance is measured. Inspection UEMs employ HCI experts and can be used in various stages of product development. Inquiry UEMs include the users' preference desires and behaviour and try to constitute the requirements of a design. According to Kjeldskov and Graham (2003), the six main areas of concern for usability testing are the: Skills and/or technological competence of the user; costs associated with data collection; the presence of researcher changing the phenomenon of interest; time and/or personnel required; inability to completely control experimental variables; observations that do not generalize.

A wide range of user-friendly UEMs involve users in activities like scenario-based tasks, interviews, focus groups. Depending upon whether the evaluation is formative or summative the size and tasks for the users vary. It is then the responsibility of the usability experts to use great judgment and interpretation skills in identifying issues.

Laboratory versus Field-Based Settings

The biggest question that the researchers are trying to find an answer to, for mobile HCI evaluations is whether to conduct the testing as part of the development process in a laboratory or in the field. Usability evaluations for mobile devices have the compounded variability in the types of users and uses of the devices along with the variability in the types of environments in which the devices can be used. According to some researchers the close connection between HCI and the computer sciences is one of the major reasons for dominance of laboratory setting for usability evaluation. As per current industry trends, nearly 70% of the usability evaluation is being conducted in laboratory settings. This poses another difficulty for the usability evaluation, in terms of developing real-life conditions in a laboratory setting.

During a laboratory usability evaluation of a mobile device or an application the testing takes place in a closed environment. Some researchers believe that laboratory evaluations provide highly valuable and quantitative data. In this environment the users are closely observed to note the way they interact with a mobile device or application to evaluate its usability. It is considered that when in a laboratory setup the users are given pre-defined tasks, the usability experiment tests all aspects of usability. Since all evaluations are generally conducted by expert technicians, in certain cases the environment can also be used to isolate users from conditions that are not relevant to the experiment. These advantages make

lab experiments useful for accurately comparing different mobile designs and interfaces.

However, isolating users from environmental factors that can affect usability may cause differences in the user experience, and the effect of environmental factors prevalent in the real world may not be felt. Also, some researchers argue that laboratory settings do not take into account environmental pressures such as noise, interruptions, movement, and multitasking. To overcome these issues they suggest a hybrid approach of using both laboratory-based and field-based evaluations.

A field study is a more generic way for collecting data about users, user needs, and product requirements. It mainly involves observation and interviews. Field study helps researchers in collecting data about task flows, inefficiencies, and the organizational and physical environments of users. Instead of trained technicians of labs, field studies rely upon investigators. The investigators observe users. While the users are involved in any activity, they take notes by asking questions. Other techniques used by the investigators are scenario-based methods, persona creation. Various aspects that can be covered through scenario-based evaluation methods are locations and setting, movement and posture, workloads, distractions and activities, devices and usages.

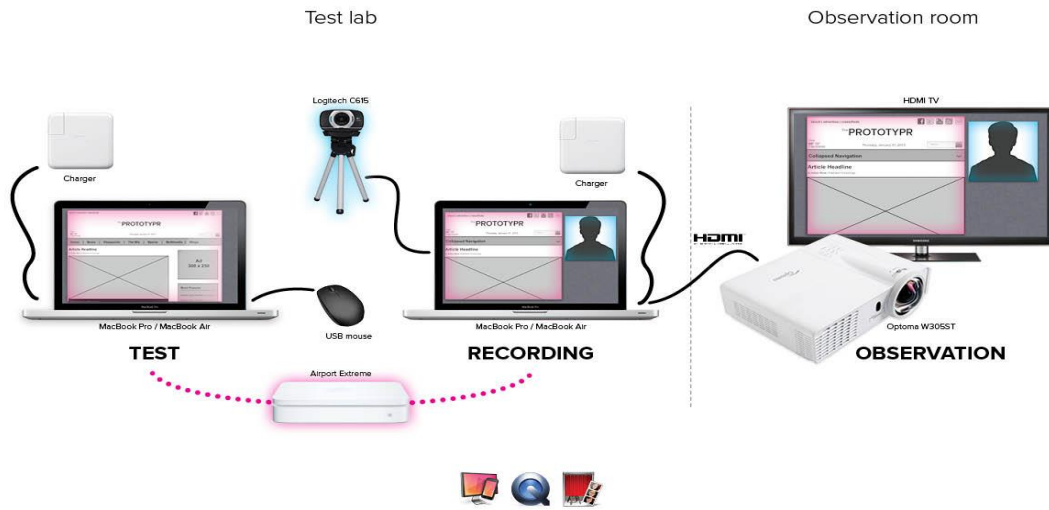
The field-based evaluations are useful in the early phases of product development as it helps gathering qualitative user requirements through the measurement of aspects such as user performance, identification of hard tasks, perceived workload, and user satisfaction. They are also. It can also be used for studying responses for already executed tasks or phases. The usability of a mobile app is measured based on participants performing tasks in a real environment. On the downside, sufficient control over users during a field study is not assured.

Here we will describe a scenario to explain the contrast in the functioning of laboratory-based and field-based evaluation for a new mobile application. Let us assume that the field tests were conducted in a university campus where the users could walk, stand still, sit or do whatever they would normally do. In the laboratory, the technician describes the defined tasks orally to the user one at a time, using the agreed wording.

So the test participants for the lab-based evaluation were supposed to visit the lab for the duration of one week for one hour every day. There they would be exposed to the new features of the application on a daily basis and the technician would observe the participants. The lab technician would then note the user experience and after a week the results of the tests would be analyzed.

Similarly in the same time frame, a group of investigators would visit the university campus and tell the participants about the application and its features. Now the users would be free to navigate around the campus and use the app as per their needs. The investigators would simply follow the participants for most of the session and note their observations. They would also note the activities and the factors around the participants. Later towards the end of the session the investigators would sit with the participants and ask them some questions about the application and its features. The information would be collected over the week and analyzed to see patterns and user response at the end of the week. Figure 2. depicts a sample lab-based evaluation setup.

Figure 2: Lab-based Evaluation Setup



Below Table 1. gives a sample laboratory based evaluation description. This table describes test scenarios and tasks descriptions for the same.

Table 1. Sample Test Scenario and Tasks involved

Task Number	Scenario of action	Task description
1.	You need to inform your friend about your personal particulars as he needs to fill up a form for you. You decide to call out.	<ol style="list-style-type: none"> 1. Dial out to contact Gerald from mobile phone contacts list. 2. Start a conversation upon pick up as you normally would. 3. Verbally inform the contact your full name, NRIC, address and date of birth
2.	You receive a call from a friend on your mobile phone. You answer the phone call.	<ol style="list-style-type: none"> 1. Answer phone call as you normally would. 2. Start a conversation with the friend.
3.	You need to inform your friend about your personal particulars information as he needs to fill up a form for you. You decided to SMS out.	<ol style="list-style-type: none"> 1. Compose a SMS including the following information: your full name, NRIC, address and date of birth. 2. Send SMS to Gerald from mobile phone contact list. 3. Reply again to Gerald if necessary, i.e. if Gerald replied your message.
4.	You received a SMS earlier and stored in the mobile phone's Inbox. You decide to reply now.	<ol style="list-style-type: none"> 1. Go to mobile phone SMS Inbox. 2. Reply to the SMS from Gerald. 3. Reply again to Gerald if necessary, i.e. if Gerald replied your message again.
5.	It is your friend's birthday and you decide to	<ol style="list-style-type: none"> 1. Create a MMS with the following

	create a Multi-media Message (MMS) and save it, to be sent out later in the day.	parameters: 10s Slide timing, Text on top, insert Image001.jpg from Images folder, and text content "Happy Birthday" 2. Preview the MMS and save it. 3. Access the saved items folder and check the content of the MMS.
--	--	---

Discussions

Once the data is collected from both the settings the analysis of this data would help us answer some basic questions. Some question that would need to be answered for such an arrangement would be:

1. What was the hypothesis for both lab-based and field-based evaluations?
2. Which out the two performed better than the hypothesis?
3. What is the effect of the environment on the participant?
4. What were the differences in usability problems found in the laboratory and the field environment when testing with a participant?
5. What were the similar problems that were observed in both the environments?
6. What was the degree of severity of the problems between laboratory and the field?
7. What was the differences in task execution times for scenarios in both environments?
8. What can be concluded for test design point of view?

The analysis of the collected data would help us identify some problems in the product design. The problems can be categorized as follows:

1. Critical problems; these problems would:
 - Prevented participants from completing tasks
 - Recur across all test scenarios
2. Serious problems; these problems would:
 - Increased test completion time severely
 - Recur frequently across test subjects and
 - Let participants still complete the task eventually
3. Cosmetic problems, these problems would:
 - Increase participants time to complete task slightly
 - Recur infrequently across test scenarios
 - Let participants complete task easily

Since both settings have their own advantages and disadvantages. Some of the possible observations

that can be made after the tests are:

1. Limitations of the lab experiments

The laboratory experiments need to be performed in controlled environments, which are aided by costly equipment. The high cost of these equipment is a primary reason for lab based evaluations to be costlier than field evaluations.. According to Duh et al. “There were many more types and occurrences of usability problems found in the field than in the laboratory. Those problems discovered tend to be critical issues regarding usability. Some of these usability problems are only related to the device being used in the field, which could not be found using conventional laboratory usability tests. With regards to the users’ behaviors, users behave less positively and more negatively in the field than in the laboratory. Some behaviors can only be observed in the field. Users also take longer time to perform certain tasks and also present more negative feelings, such as dissatisfaction and difficult of use, to the use of the device in the field.”

2. Updates mobile user interface

The evaluations can be performed for any mobile device or a mobile application. Under all circumstances it is suggested that the evaluations be done on the latest hardware and software specifications. As in the modern time all devices keep updating, the product must be designed in such a way that it supports those updates and performs with the same level of competence.

3. Usability measures

The usability evaluators need to pay attention to detail in all phases of the product development. They should clearly state their expectations to the participants. The evaluation environments should have clearly documented rules for documenting the user feedback.

4. Measurement methods

The usability evaluators need to establish standard measurement methods for all environments. They should have well documented rules for lab-based evaluations. Since data collection in field-based evaluations is dynamic the evaluators must have the scope to accommodate any new information. They should also define the ways of processing any unexpected data which is captured during the evaluations.

Conclusion

Once the observations have been documented and analyzed the evaluators should measure the results from both environments and use the best results for developing their product.

References:

- Klockar, T., Carr D. A., Hedman, A., Johansson, T., & Bengtsson, F. (2003). Usability of mobile phones. *Proceedings of the 19th International Symposium on Human Factors in Telecommunication*, 197–204.
- Grudin, J. 2009. “Brian Shackel’s contribution to the written history of Human–Computer Interaction,” *Interacting with Computers* (21:5), pp 370-374.
- Jaimes, A. and Sebe, N. 2007. “Multimodal human-computer interaction: A survey,” *Computer Vision and Image Understanding* (108:1-2), pp 116-134.
- Blandford, A. 2004. “*Understanding Users’ Experiences: Evaluation of Digital Libraries*,” in *Proceedings of the Delos workshop on evaluation of Digital libraries*, Padova, Italy, pp 31-34.
- Kjeldskov J. and Stage J. (2004) *New Techniques for Usability Evaluation of Mobile Systems*. *International Journal of Human-Computer Studies (IJHCS)* Elsevier (forthcoming 2004)
- Kjeldskov, J and Graham, C. (2003) *A review of MobileHCI research methods*. Proc MobileHCI 2003.
- Rubin J. *Handbook of Usability Testing*. Wiley, New York. 1994.
- Law, E. L., & Hvannberg, E. T. (2002). Complementarity and convergence of heuristic evaluation and usability test: *A case study of UNIVERSAL brokerage platform*. *Proceedings of the Second Nordic Conference on Human-Computer Interaction* (pp. 71–80). New York: ACM Press.