Usability Evaluation Approaches for (Ubiquitous) Mobile Applications: A Systematic Mapping Study

[1]Rodrigo A. Cruz Reis, [2]Ludymilla L. A. Gomes, [3]Arilo Claudio Dias-Neto, [4]Awdren de L. Fontão

[123]Institute of Computing

Federal University of Amazonas

[4]INDT

Manaus-AM, Brazil

Emails: [1]rdgdosanjos@gmail.com, [2][llag] [3][arilo] @icomp.ufam.edu.br, [4]fontao@microsoft.com

Abstract—In case of ubiquitous mobile applications, there is an increased need for effective/efficient approaches to evaluate the usability of these applications. The technical literature provides several evaluation approaches found in several sources, with different characteristics and classifications. This paper presents the results of a systematic mapping study that investigated usability evaluation approaches for (ubiquitous) mobile applications. In total, we identified 101 usability evaluation approaches for mobile applications, 28 of which applied to ubiquitous mobile applications. They were classified according to some attributes, such as: type of evaluation technique, type of mobile apps to be evaluated, experiment used to evaluate the approach, usability attribute/factors to be evaluated, and characteristics of ubiquity evaluated by each approach, representing the state-of-art in this research field.

Keywords-Usability; Mobile apps; Systematic Mapping; Survey.

I. INTRODUCTION

Smartphones have become very popular in our current society. Advances in mobile technologies have allowed the emergence/development of a wide range of software for these mobile devices (called mobile applications or, simply, mobile apps) [1]. This platform introduced several advantages. Perhaps, the most noticeable would be the mobility to its users while using different mobile apps.

This large and growing number of mobile apps has challenged software engineers to develop applications with a high level of quality in order to become more attractive and competitive in this new market [2]. Moreover, this platform introduced some challenges and constraints to be considered during the software development, such as small screen size, limited connectivity, high power consumption rates and limited input modalities [10].

According to Duh et al. [3], usability is a critical factor for the popularity and success of mobile apps. A good usability design improves the device user's operability and, thus, enhances the overall product quality. Users tend to choose applications that are easy to learn, which take less time to complete a particular task and seem to be more "friendly" to the user [4]. Thus, various approaches aimed at supporting the usability definition and evaluation for mobile apps have been proposed in the technical literature.

Usability evaluation of software for desktop and mobile devices platforms is an emerging areas of research [5]. In the past, software usability was subjectively evaluated by informal processes [4]. Researchers just selected some usability attributes that they wanted to assess and then measured what they considered important. In recent years, usability measurement and analysis approaches have been proposed and improved. Laboratory experiments, field studies and hands-on measurements are some of the methods most often applied by researchers [4][6]. Each of these evaluation methods has its advantages and disadvantages. Due to the highly dynamic context of use, offered by mobile apps, laboratory and field usability testing involves different challenges and may find different usability problems [3].

In order to analyze the scenario of evaluation usability approaches for ubiquitous mobile apps and ubiquitous mobile apps, this paper presents the results of a systematic mapping study [7] that identified and characterized 101 different approaches. This study aims to complement previous characterization studies, such as the studies published in [4][6], in two aspects: (1) it updates the list of approaches identified in the technical literature; (2) it presents a different perspective on the identified approaches, analyzing, for example, the categories of mobile apps and the type of proposed evaluation approach (e.g., static or dynamic analysis). Finally, some challenges and trends are presented as a result of this study.

This paper is structured as follows. Section 2 presents some definitions relevant to this paper and related work. Section 3 presents the systematic mapping protocol. Section 4 presents the results obtained. Finally, Section 5 presents a summary of this work and a brief discussion on future work.

II. USABILITY EVALUATION IN MOBILE APPS

Several definitions for usability can be found in the technical literature. For example, ISO-9241 [8] defines usability as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specific context of use".

With the emergence and rapid deployment of mobile technologies, the usability of applications developed for this platform has been the focus of several studies. According to Duh et al. [3], a good project, besides meeting the needs of the market and providing the device user satisfaction, can also reduce physical and mental stress, reduce the learning curve, improve the operability of device use and, thus, improve the overall product quality.

Zhang et al. [10] claimed that the mobile usability includes some of the new challenges related to mobility, such as: mobile context, connectivity, small screen size, different display resolutions, limited capacity and power processing, data entry methods, interaction with multi-touch screen,

show different resolutions and dimensions, device orientation changes and gestures such as tap, flick, and pinch. Thus, an approach aiming to evaluate the usability of mobile apps needs to deal with these challenges.

Software usability evaluation approaches have become increasingly popular in technical literature [11],[12]. Usability evaluation approaches aims to obtain a third-party judgment regarding user's characteristics to assess effectively and efficiently whether a user is able to view the content or perform a task on a specified device [10].

Some previous studies presented a usability evaluation analysis for mobile apps. In [4], the authors presented a study that analyzed the methodologies used to empirically evaluate the mobile usability, classifying as laboratory experiments, field studies and measuring practice. The study described advantages and limitations of each method, but did not identify/characterize the publications on usability evaluation in the technical literature.

In [6], the authors presented a longitudinal review of Human-Computer Interaction (HCI) research methods for the mobile platform published until 2012, analyzing more than 140 papers. In this study, publications were classified in terms of their research method (case study, field of study, action research, laboratory experiment, survey, basic and applied research, and normative writings) and purpose (understanding, engineering, re-engineering, evaluation and description). This study revealed that 68% of the material evaluated in research on human-computer interaction in mobile apps until 2009 involved mobile usability evaluations, where 63% of these researches made through laboratory experiments, 29% through field studies, and 7% through surveys.

Duh et al. [3] described a study that investigated the differences between the usability testing on mobile phones conducted in laboratory and real-life situations. Significant differences were found, including the frequency and severity of usability problems found in both scenarios, user behavior and subjective responses to the device and the interaction between users and the devices.

Kjeldskov et al. [5] presented and analyzed six techniques for evaluating the usability of mobile apps in laboratory. The six techniques were analyzed using two usability experiments. The goal was to examine whether the evaluation of mobile systems in a controlled environment is similar to a real user behavior.

Finally, in [10], the authors presented an overview of existing usability studies, focusing on usability testing, and discussed the main issues investigated in the technical literature. Then, they proposed a generic framework and provided detailed guidelines on how to conduct such usability studies.

We can observe that the studies and approaches that address the evaluation of usability in mobile apps are dispersed in different sources in the technical literature, making it difficult to analyze empirical evidence known about this research area. Thus, this paper describes a systematic mapping study conducted to identify/characterize/evaluate usability evaluation approaches for mobile applications proposed at the technical literature. The following sections present the planning and results of this systematic mapping study that investigated different perspectives related to usability evaluation approaches for mobile apps.

III. SYSTEMATIC MAPPING ON USABILITY EVALUATION FOR MOBILE APPS

According to Kitchenham et al. [7], a systematic mapping consists of a type of secondary study where the dimensions to be evaluated in a secondary study (population, intervention, comparison and outcomes) are not fully described. This study explores a less strict research protocol when compared to protocols commonly used in systematic reviews.

A good systematic mapping always considers the following questions [13]: identifying all published materials related to the investigation goal; choosing criteria for the inclusion of materials; evaluating the quality of each material; producing the results of each material impartially; interpreting the results; and, presenting a reasonable and neutral summary of the results.

This research follows a systematic mapping process described by [7], which is composed of three stages: (1) Plan the study; (2) Conduct the review; (3) Report the results. The activities related to the planning and conducting of this literature systematic mapping study will be described in the following subsections. The results from this study are described in the subsequent section.

A. Research Questions

The objective of this study is to identify approaches and types of research in usability evaluation for mobile apps and also point out the areas where the available empirical evidences were insufficient and therefore, more studies are needed. In order to address the objectives of this research, four relevant research questions were prepared:

- Q1. What types of approaches have been proposed for usability evaluation of mobile apps?
- Q2. To which category of mobile app approaches are employed usability evaluation approaches?
- Q3. What usability attributes/factors are evaluated by these approaches?
- Q4. Which characteristics of computational ubiquity are evaluated by this approach?

B. Identifying and Selection of Primary Studies

The sources used for selection of primary studies in this study were two digital libraries: IEEExplorer and Scopus (according to its maintainer, this online indexing service would cover the major computing digital libraries, such as ACM Digital Library or Science Direct. Only IEEExplorer would be partially indexed by Scopus).

The search string used for the search of primary studies was structured according to the rules described in [14], and was composed of the elements Population (P), Intervention (I), Comparison [optional in a systematic mapping study] (C) and Outcomes (O), as follows:

- **Population:** "Mobile Application" OR "Mobile Software" OR "Mobile App" OR "Mobile System".

- **Intervention:** "usability" OR "user experience" OR "HCI" OR "human computer interaction".
- **Comparison:** not applied to systematic mapping study.
- **Outcome:** "evaluation" OR "assessment" OR "measure" OR "experiment" OR "test" OR "inspect" OR "review".

C. Primary Studies Inclusion Criteria

A list of primary studies was obtained through the search string from the selected sources of bibliographic material. Then, the following criteria for inclusion of primary studies that were related to the objective of this study, in order to answer the research questions, were applied: (1) It describes research that explores usability evaluation approaches for mobile apps; (2) It must contain a full research publication; (3) It must be written in English; (4) It must be available for download.

Papers duplicated on different search sources (e.g., papers indexed by IEEExplorer and Scopus) would have only one instance selected in this study.

D. Systematic Mapping Execution

The activities of execution of research string and papers selected in this study were made between January and February 2014.

TABLE I.   NUMBER OF SELECTERD PUBLICATIONS PER PHASE

Source	Returned Papers	Filter 1	Filter 2	Filter 3
IEEEXplorer	317	10	8	1
Scopus	53	170	93	27
TOTAL	**370**	**180**	**101**	**28**

The preliminary research offered 317 relevant publications in the Scopus library and 53 in the IEEExplorer library. The inclusion analysis of these papers was done in three steps (Table I):

1. Tracking the initial set of papers based on the titles, abstracts and introduction sections. In total, 180 publications were pre-selected to step 2;
2. Complete reading of the paper. A total of 101 publications were selected. In this step, the information to answer questions Q1, Q2 and Q3, previously presented in section III.A, was extracted from these papers.
3. Complete reading of the article from the ubiquitous applications point of view. 28 publications were selected in this phase, which were used to answer Q4.

E. Data Extraction Form

For each selected paper, we extract the main information aiming to characterize the usability evaluation approach:
- **[YEAR]** Publication Year.
- **[CATEGORY]** Type of evaluation technique:
  - Static: methods that do not involve software execution.
  - Dynamic: methods involving software execution with real/simulated data in a real or simulated environment.
- **[TYPE]** Type of mobile apps evaluated by the approach, classified according to [14] as:
  - **Native Apps:** application specifically developed to execute on a specific device platform.

- **Web Apps:** application that runs over a browser embedded in the device and does not have access to some device's internal resources.
- **Hybrid Apps (HTML5 and widgets apps):** they get stored in the device's main screen and can take advantage of all devices' internal resources, but they can be based on HTML5 and displayed through a web browser.
- **[EVALUATION]** Type of empirical evaluation applied to the approach, according to [6]: Case Study, Field Study, Action research, Lab experiments, Survey research, Applied research, Basic research, Normative writings.
- **[ATTRIBUTES]** Attributes evaluated by the approach (classification proposed by [15]): Efficiency, Satisfaction, Learnability, Memorability, Errors.
- **[FACTORS]** Usability factors evaluated by the approach (rating also proposed by [1]):
  - **User:** It is important to consider the end user of an application during the development process.
  - **Task:** refers here to the goal the user is trying to accomplish with the mobile application.
  - **Context of Use**: refers here to the environment in which the user will use the application.
- **[UBIQUITOUS]** Ubiquitous characteristics evaluated by the approach (classification proposed by [16]): Pervasiveness services, Invisibility, Context awareness, Adaptive behavior, Experiences Capture, Functionality composition, Spontaneous interoperability, Heterogeneity of devices, and Fault tolerance.

IV. RESULTS ANALYSIS

The usability evaluation approaches for mobile applications were analyzed according to the characteristics defined in the data extraction form (Section III.E). Thus, the overall results for each research question (presented in section III.A) and attributes extracted from the evaluation approaches are discussed in subsequent sections.

A. Analysis by Publication Year

In this study, we identified usability evaluation approaches for mobile apps published from 2004 to 2014 (when the study was run).
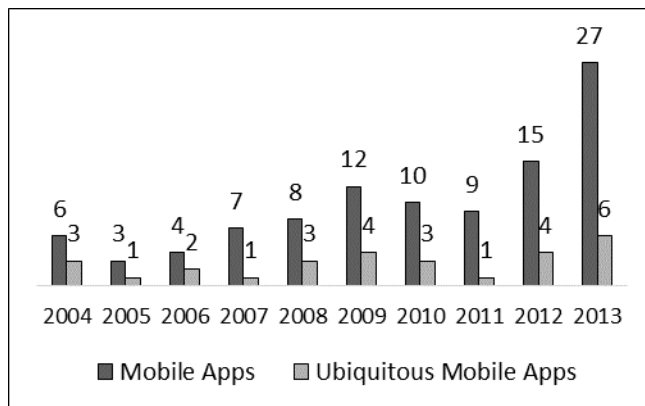

Figure 1.   Analysis of papers by publication year.

The distribution of results is displayed graphically in Figure 1, where it is observed that there was a considerable increase in research on usability evaluation for mobile apps in the community recently (mainly in the past 2 years). This indicates the need for research in this area and shows the evolution in the level of importance of the issue. With the advances in mobile technology in bringing the concept of ubiquity, this area tends to become more interesting for future research [15].

We also noticed the number of usability evaluation approaches for ubiquitous mobile apps remained stable from 2009 until 2013. This 5-year interval has 18 of the 28 papers found in this study. This shows that the need for evaluation in ubiquitous mobile apps is really relevant to the academy and its interest in academic research is growing in the last years.

B. Analysis by Type of Evaluation Technique

In order to answer the question Q1 discussed in section III.A, an analysis of evaluation techniques per category was made (Figure 2).
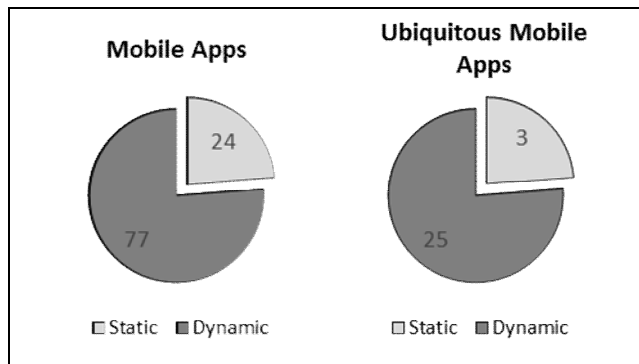

Figure 2. Analysis by evaluation technique category

We observed the category of usability evaluation techniques most frequent are dynamic approaches (77/101 for mobile apps; 25/28 for ubiquitous mobile apps). This result is justified due to the need of assessing the app on a scenario closer to reality, possibly by using dynamic approaches, making the evaluation more efficient.

C. Analysis by Mobile App Category

In order to answer the question Q2 discussed in section III.A, an analysis per mobile app category evaluated by the identified approaches was performed (Figure 3).

Among the categories analyzed in this study, it is remarkable that native apps have been more explored in research with the purpose of evaluating usability attributes in general and ubiquitous mobile apps. The reason could be the requirements, accessibility, and restriction issues imposed by mobile platforms. Then, the second more explored category is web apps, due to the popularity of this type of application for the mobile platform.
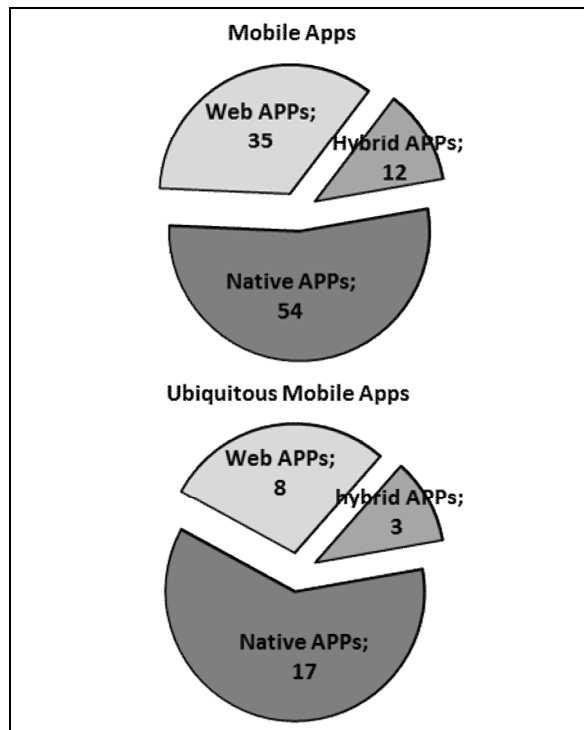

Figure 3. Analysis of mobile application category

Finally, evaluations in hybrid apps are starting to emerge, because it represents a new trend of development of mobile apps, justifying the small number of research in the area. All authors who proposed usability evaluation approaches for hybrid apps highlighted that this category is emerging and needs more research, not only for evaluation of usability, but also for application development.

D. Analysis by Empirical Evaluation Type

We also analyzed the type of empirical evaluation applied to the selected approaches, as shown in Figure 4. The results indicate that several authors have chosen to apply empirical techniques as a strategy for the final assessment of the proposed approaches. The results indicate the predominance of Case Studies (40/101), followed by Field Studies (29/101) and Lab Experiments (25/101). Three type of empirical evaluation were not found in the selected papers: Normative writings, Applied Research, and Basic Research. Analyzing the results for Ubiquitous Mobile Apps, they indicate approximate values between the same three types of evaluation: Field Study, with (10/28, followed by Case Study (9/28) and Lab Experiment (8/28).
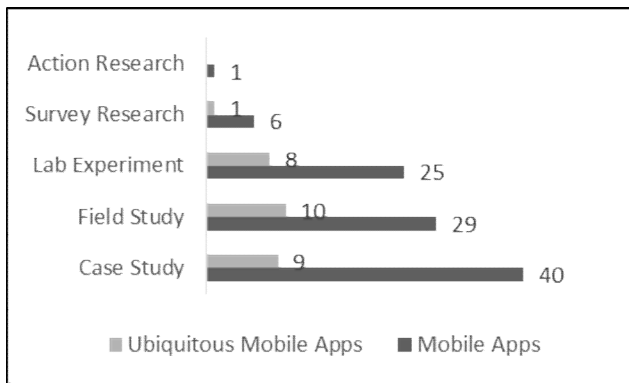
Figure 4.   Analysis by type of experiment.

Trying to understand the result, we could observe the types of investigation most applied to evaluate the proposed approaches (case, field, and lab study) were formal investigations, having more credibility in the academy. This scenario can justify the difference obtained when compared to other types of investigations.

E. Analysis by Usability Attributes

In order to answer the question Q3 addressed in Section III.A, an analysis by usability attributes was performed, as shown in Figure 5. We observed that user satisfaction is the most investigated attribute in the identified approaches (67/101 papers) on mobile apps, and it is the second more investigated in ubiquitous mobile apps (15/28 papers). In this analysis, a paper could address one or more attributes, justifying that the sum of the numbers distributed among the attributes is greater than the number of identified papers.
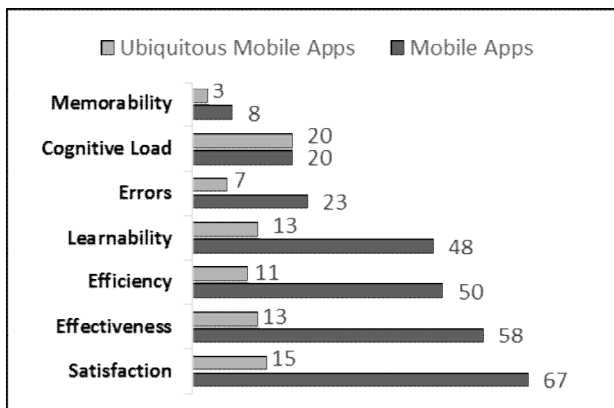

Figure 5.   Usability Attributes Analysis

It is not possible to conclude that a usability attribute, for being more addressed than others, would be more or less relevant to be evaluated for mobile apps. The results only describe the state of the art on the use of these attributes in the academy. However, an interesting aspect may be observed: the Cognitive Load attribute was only used in mobile apps that deal with context awareness requirements, one of the features present in ubiquitous mobile apps (question 4, to be discussed below).

F. Analysis by Usability Factors

Yet to answer the question Q3 addressed in section III.A, an analysis by usability factors was done, as shown in Figure 6. Analyzing the results, we observed that the factors user (60/108) and tasks (66/108) are more frequent in the selected papers. In general, most of the papers when dealing with one of these factors also deal with the second one. The evaluation of the factor context of use was observed only in approaches that deal with context awareness requirements. In this analysis, a paper could address one or more factors, justifying that the sum of the numbers distributed among the factors is greater than the number of identified papers.
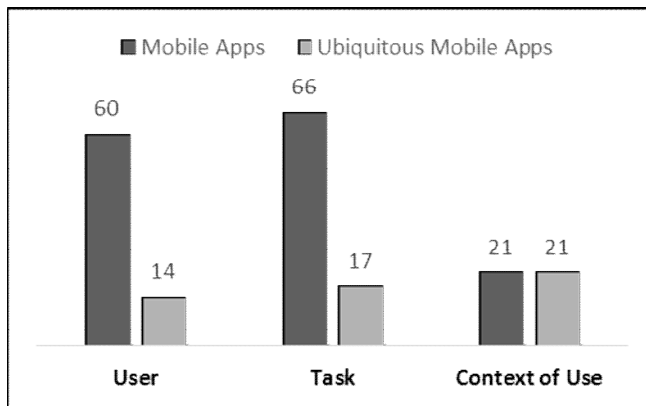

Figure 6.   Analysis of Usability Factors

Furthermore, this also shows us that ubiquitous mobile applications can be evaluated recitals three factors as context of use with 21 identified papers, the user featuring 14 papers and 17 papers task with a total of 28.

G. Ubiquitous Features Analysis

In order to answer the question Q4 addressed in Section III.A, an analysis of ubiquitous feature evaluated by the identified approaches was done (Figure 7). In [16], a table with ubiquity characteristics from a functional point of view is presented.

Only 28 papers addressed the usability evaluation for mobile apps with characteristics of computational ubiquity. Among the 10 characteristics defined in [16], only 5 were addressed in papers identified in this study, suggesting that these would be the computational ubiquitous features that could be evaluated by means of usability requirements.

Furthermore, we observed that the definition of pervasiveness services suggests it as a main feature of ubiquity. Thus, in the case of ubiquitous mobile apps, pervasiveness services feature will always be present. This explains why all 28 papers that deal with usability evaluation approaches for ubiquitous mobile apps cite this feature, but did not propose an approach to analyze this feature for ubiquitous mobile apps.

The distribution of the papers among the ubiquitous characteristics is presented in Figure 7.
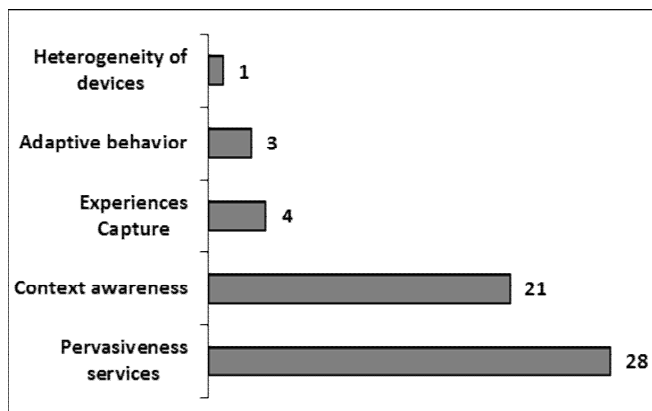
Figure 7. Feature of ubiquity analysis

By analyzing the results, we observed that the characteristics context awareness was addressed in 21 papers. We observed that in other characteristics, few studies addressing usability evaluation were found. One factor that may hinder this analysis is that each ubiquity feature has a vast array of settings and areas that still need to be analyzed from a usability point of view. We could observe that research in the field of usability in mobile apps that deal with each ubiquity feature is still quite scarce. This shows good opportunities for research in this domain area.

V. CONCLUSIONS

The number of mobile apps used in daily life is continuously growing and so is the search on their quality. Despite this evolution, if we compare the demand implementation of ubiquity characteristics, which is a factor present currently in many apps, we can observe that there is still a need for studies on mobile apps usability evaluation.

The results of this systematic mapping study revealed some perspectives about the approaches to support the evaluation of usability in mobile apps in the last 10 years. For example, they indicate that such approaches mainly utilize dynamic techniques (e.g., testing). Many publications brought justification for use of such technique, saying that the context of use was the main reason for choosing the testing technique.

We also observed that native and web apps have been the focus of usability evaluation approaches for mobile apps, which indicates a need for attention to hybrid apps, which are partly native and partly web application.

In order to evaluate the techniques, experimentation is ahead of the other techniques. The justifications of the authors are related to the restrictions that mobile apps need, what could be settled in an empirical evaluation.

We also observed that a small number of studies covers ubiquitous mobile apps. Soon, there will be the necessity for more studies on this topic.

This systematic mapping was done in order to identify which types of research in mobile app usability evaluation are being used in the academy. With the results obtained from the mapping, it is possible   identify the areas most addressed by the community in which there are a large number of studies and point out the areas where the available evidence is insufficient and therefore more studies are needed.

The need for studies focused on different ubiquity factors oriented mobile apps is noticeable. Future work can be made, such as: choose the type of mobile application category that can be web applications, native applications or hybrid applications and instantiate ubiquity factors to usability evaluations. There is a clear need for approaches, processes, tools to support the assessment of usability in ubiquitous mobile applications.

REFERENCES

[1] R. Harrison, D. Flood, and D. Duce, Usability of mobile applications: literature review and rationale for a new usability model, "Journal of Interaction Science 1 (1), 2013, pp. 1-16.

[2] Global mobile statistics 2011, <http://mobithinking.com/mobilemarketing-tools/latest-mobile-stats> 2014/05/13

[3] H. Duh, G. Tan, and V. Chen, "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 (MobileHCI), 2006, pp. 181-186.

[4] F. Nayebi, J. M. Desharnais and A. Abran, "The State of the Art of Mobile Application Usability Evaluation," In: 25th IEEE Canadian Conference on Electrical Computer Engineering CCECE, 2012, pp.1-4.

[5] J. Kjeldskov and J. Stage, "New techniques for usability evaluation of mobile systems," In: International Journal of Human-Computer Studies, 2004, pp. 599-620.

[6] J. Kjeldskov and J. Paay, "A longitudinal review of Mobile HCI research methods," In: International Conference on Human-computer interaction with mobile devices and services (MobileHCI), 2012, pp. 69-78.

[7] B. Kitchenham and S. Charters, "Guidelines for performing Systematic Literature Reviews" in Software Engineering, EBSE Technical Report EBSE-2007-001, Department of Computer Science, University of Durham, 2007.

[8] ISO 9241-11 "Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs)" International Standards Organisation, Geneva, 1997.

[9] ISO/IEC 25000." Software Engineering -- Software product Quality Requirements and Evaluation (SQuaRE) -- Guide to SQuaRE", (2003/2014).

[10] D. Zhang and B. Adipat, "Challenges, methodologies, and issues in the usability testing of mobile applications," International Journal of Human-Computer Interaction, vol. 18, no. 3, 2005, pp. 293-308.

[11] M. Billi, L. Burzagli, T. Catarci, G. Santucci, E. Bertini, F. Gabbanini, and E. Palchetti, "A unified methodology for the evaluation of accessibility and usability of mobile applications," Universal Access in the Information Society (9:4), 2010, pp 337-356.

[12] K. Hornbæk, "Dogmas in the assessment of usability evaluation methods," Behaviour & Information Technology (29:1), 2010, pp 97-111.

[13] E. Mendes, "A systematic review of web engineering research", Proceedings of the International Symposium on

Empirical Software Engineering (ISESE), Australia, 2005, pp. 498-507.

[14] M. Pai, M. McCulloch, J. D. Gorman, N. Pai, W. Enanoria, G. Kennedy and J. M. Colford Jr, "Systematic reviews and meta-analyses: an illustrated,' step-by-step guide. Natl Med J India. Mar-Apr, 2004, pp 86–95.

[15] J. Nielsen, "Usability engineering. Morgan Kaufmann Pub," (1994).

[16] R. O. Spínola, H. T. Guilherme, "Towards a framework to characterize ubiquitous software projects, Information and Software Technology," Volume 54, Issue 7, July 2012, pp 759-785.