

ROLE OF CONTEXT IN USABILITY EVALUATIONS: A REVIEW

Munesh Chandra Trivedi¹ and Mohammadi Akheela Khanum²

¹DIT School of Engineering, Greater Noida, India
munesh.trivedi@gmail.com

²PAHER University, Udaipur, Rajasthan, India
akheela.khanum@gmail.com

ABSTRACT

Usability is often defined as the ability of a system to carry out specific tasks by specific users in a specific context. Usability evaluation involves testing the system for its expected usability. Usability testing is performed in natural environment (field) or artificial environment (laboratory). The result of usability evaluation is affected by the environment in which it is carried out. Previous studies have focused only on the physical environment (lab and field) effect on the results but rarely focused on the effect of social environment (people present during testing). Therefore, this study aims to review how important it is to take context into account during usability evaluation. Context is explored through the theory of behaviour settings, according to which behaviour of individuals is strongly influenced by the physical as well as the social environment in which they function. The result of this review indicates that the physical and social context plays a substantial role in usability evaluations. Further, it also suggests that the usability evaluation model should encompass context as an important component in the framework.

KEYWORDS

Usability evaluation, context, physical context, social context, theory of behaviour settings

1. INTRODUCTION

Now a days when a users buy any gadget, be it a mobile phone, laptop, or an ipad, they first check how easy and understandable the gadget functionality is. Whether they can perform the required operations quickly, are they able to understand the icons on the interface easily without any help. In short, the user is concentrating on the usability of the device. Therefore, usability has become an important parameter today. Usability engineering consists of a set of activities that take place throughout the lifecycle of the product. However, significant activities happen at the early stages before the user interface has been designed [1]. “Usability Engineering” is a science which is concerned with how to understand and systematically address the usability demand of a user. Thus, usability engineering deals with design of Web sites, computer portals, computer keyboard design, car dashboard design, TV remote key layouts, washing machine front panel layout, etc [2].

1.1 Usability

Usability is most often defined as the ease of use and acceptability of a system for a particular class of users carrying out specific tasks in a specific environment. Ease of use affects the user’s performance and their satisfaction, while acceptability affects whether the product is used [3]. Nielsen [4] considers that the usability of a system can have five quality components:

- (i) Learnability: how easy is it for the users to accomplish basic tasks the first time they encounter the design?
- (ii) Efficiency: once users have learned the design, how quickly can they perform tasks?

(iii) Memorability: when users return to the design after a period of not using it, how easily can they reestablish proficiency?

(iv) Errors: how many errors do users make, how severe are these errors, and how easily can they recover from the errors?

(v) Satisfaction: how pleasant is it to use the design?

Usability has been defined by the International Standards Organization (ISO) as “the extent to which the product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use” [5].

1.2 Usability Evaluation

Usability evaluation is an important part of today’s software development process as it can help improve the usability of systems under development. When introduced into the process at the right time, usability evaluations can be cost effective in terms of time and money [1]. The basic aim of usability evaluation is to improve the usability of products. Through usability evaluation possible weaknesses with regards to a system’s usability with the involvement of actual users can be identified. Usability evaluation involves presenting the users with some tasks which are reflective of the future system use. The results of a usability evaluation can be represented in different forms, such as error rates, time taken to complete the task, and number of usability problems found. Usability evaluation is generally carried out in usability laboratories (in-vitro) and in some cases can be carried out in field (in-situ). Holzinger [6] divided the usability evaluation techniques into inspection methods (without end users) and test methods (with end users). Table 1 depicts this division.

Table 1: Categorization of usability evaluation techniques

Source: [6]

Inspection Methods	Test Methods
Heuristic Evaluation	Thinking Aloud
Cognitive Walkthrough	Field Observation
Action Analysis	Questionnaires

Heuristic evaluations are expert evaluations of products or systems, including information systems and documentation. They’re conducted by usability specialists, domain experts, or preferably by “double experts” with both usability and domain experience [7]. Advantage of evaluation is that it can produce results in a limited time because it does not involve time-consuming participant recruiting. The disadvantage is that the results of heuristic evaluation cannot be fully trusted as no real users are involved.

Cognitive walkthrough is a task-oriented method by which the analyst explores the system’s functionalities; that is, it simulates step-by-step user behavior for a given task. It traces the cognitive issues, such as learnability of the user, by analyzing their mental processes. Cognitive walkthrough is known for its benefits such as low cost and quick results. It is helpful in picking out interface problems at an early stage. However, it can be sometimes time-consuming, and since restructuring the interface is often expensive and difficult at later stages in development; the cognitive walkthrough is usually applied in early stages of software development.

Action analysis involves an inspection of the user actions with regard to physical, cognitive, and perceptual loading. It is helpful in predicting the time a given task will take to complete and also

helps in getting an insight into the users' behavior. The short coming of action analysis is that it requires high expertise and time.

With the think aloud protocol, a user is required to verbalize his comments about areas where he is struggling and the reasons for the difficulties. The outcomes of think aloud protocols can be used by the usability practitioner to identify problem areas of the Web site or application being assessed and to find appropriate improvements. One of the most common think aloud protocols that usability practitioners engage in today is concurrent think aloud under which the participant is encouraged to "think out loud" while working on a task [8].

Field observation is the simplest of all methods. It is carried out by visiting the user's workplaces and observing their use with target interface or system. Possibly, notes about the major usability problems being faced by the user are taken by the observer.

Questionnaires are indirect usability measures. They don't study the interface directly; rather collect the user's view about the interface. Questionnaires have to be designed by the experts and should cover all the experiences with the interface. In order to validate the results of the questionnaires large number of users has to be assessed.

Usability evaluations cannot be simply based on the results of application of one or more of the above techniques. Many aspects of context, such as the users, the location, and the culture, all of which can be important during the evaluations, have to be taken into account. Therefore, in what follows we explore what context is? What aspects form the context? And what is the role played by the context in usability evaluations?

2. MOTIVATION

In order to understand the relation between the context and the individual's behaviour we chose the widely accepted theory of behaviour settings. The theory of behaviour settings was introduced by Roger Garlock Barker in late 1940s [9]. He continuously collected empirical data from a small town in Kansas with less than 2000 people from 1947 through 1972 based on which he developed the theory of behaviour settings. Behaviour setting theory proposes that there are specific, identifiable units of the environment, the physical and social elements, which are combined into one unit, which have very powerful influences on human behaviour [10]. A behaviour setting is a naturally occurring unit of the environment at the molar (perceived as wholes as opposed to parts) level, recognized by its inhabitants, that is, people perceive that they conduct their lives inside behaviour settings [11;12;9]. Barker [11] found that psychology involves both the psychological and the ecological environment. He examined that the distinction between human psychology and his ecological environment was difficult. Therefore, Barker focused on molar human behaviour rather than individual units. For example, he interpreted the act of buying a stamp as an entire unified behaviour not broken down into micro acts that followed the stamp buyer through the myriad of smaller components of the total act [9]. The interface of the ecological and molar behaviour creates ecological units [11]. According to Schoggen [9] these units arise simultaneously in physical, social, psychological and behavioural realms and share three common attributes:

1. They are self-generated, independent from the observer's or researcher's interest or manipulation.
2. They have a time/space locus.
3. They have a boundary which separates the internal pattern of the unit from the external pattern of the surround.

A behaviour setting is a pattern of ecological units and consists of "standing pattern of behaviour" Barker [11]. Barker described the standing pattern as a milieu (settings), circumjacent, and synomorphic or fitting to the behaviour [9]. The behaviour is happening in a milieu and milieu matches the behaviour.

The close interrelation of location/settings and people as seen through the Barker's theory of behavior settings could be an indication that context is important and plays a vital role in influencing the results of usability evaluations.

3. CONTEXT IN USABILITY EVALUATION

Context is a term defined differently by different people. For example, Brown et al. [13] define context as location, identities of the people around the user, the time of the day, season, and temperature. Ryan et al [14] define context as the user's location, environment, identity and time. Hull et al [15] defines context to include the entire environment aspects of the current situation. Schilit et al [16] claim that the important aspects of the context are: where you are, who you are with, and what resources are nearby. Dey et al [17] define context to be the user's physical, social, emotional or informational state.

Taking into consideration the various definitions of context, we found that one of the likely definitions of the context would be "*context is anything which has an effect on the human behaviour*". When evaluating the usability of any system, the behaviour of the user is very important. The factors which may affect the user behaviour needs to carefully considered because the result of usability evaluations may vary in different settings where the user may exhibit varying behaviours.

Product usability is not an independent activity but depends on the context. It is not meaningful to talk simply about the usability of a product, as usability is a function of the context in which the product is used. The characteristics of the context (the users, tasks, and environment) may be as important in determining usability as the characteristics of the product itself. Changing any relevant aspect of the context of use may change the usability of the product [19]. To understand the role of context in usability evaluation, it is necessary to examine what context is, and what aspects it comprise of. Many studies report that just physical location is the context. Context can be the cultural context [20], organizational context, technological context or social context [21]. Since our focus is on the physical and social context, we explore each one of them in the next two sections.

3.1 Physical Context

Physical context comprise of the physical surroundings of the users, it is the location, the place where usability evaluation takes place. Physical context usually refers to the environment in which user is tested. Natural environment is the location of actual use of the system being tested. Usability evaluation taking place in natural environment is called as the field testing. Artificial environment is the simulation of natural environment, sometimes referred to as the controlled environment. Usability evaluation carried out in artificial environment is also referred to as laboratory testing.

Traditionally laboratory experiments are employed to evaluate the usability of computer systems, and to improve the understanding of usability. Laboratory testing takes place in a controlled environment with the experimenter in control of assignments of subjects, treatment variables and manipulation of variables. It is possible to employ facilities for collection of high-quality data such as video recording of the display and user interaction [22]. Razak et al. [23] states that the evaluation done in the laboratory has several advantages. First, the conditions for conducting research can be controlled. Secondly, all the participants experience same setting leading to higher quality data. Laboratory studies allow the researchers to focus on specific phenomena of interest and facilitate good data collection. Laboratory testing has received both appreciation as well as criticism. The laboratory evaluations do not simulate the context when usability testing is done with mobile phones, because laboratory settings lack the desired ecological validity [24]. Even though the adequateness of laboratory evaluations is questioned, 71% of the mobile device evaluations were done in laboratory settings [25]. Similar claim also comes from Park & Lim [26] where they state that simulating the use settings is very hard , time consuming, expensive and lacks contextual factors.

Field testing takes place in a more natural setting. An artificial setting supports control but

lacks realism whereas a natural setting supplies realism but makes control more difficult [27]. Oh & Kim [28] claims that to test children requirements and issues with the everyday use of technology, field evaluation would be a better choice. This claim is also supported by Xu et al. [28] in search for evaluation methods for children's tangible technology where they found that location plays a large part in how children behave, children felt more at ease and focused when they are at school.

3.2 Social Context

A major part of the context of a usability evaluation is the people involved. This is also often referred to as the social context of the usability evaluation. People involved in usability evaluations can be the test evaluators, the test monitors, the users. There also exists a substantial effect of the other people who may not be directly involved with the evaluation. Such people can be the parents and teachers of the children users and the onlookers. Stoica et al. [30] found that while laboratory evaluations give excellent data, the context and the surroundings as well as other people around also play an important role. One of the purposes of creating a social context in usability evaluations is to facilitate effective and efficient evaluations. Creating a proper social context can potentially diminish some of these challenges [31]. Although social context is considered important, only little research has been done to identify how it influences usability evaluations.

4. RELATED WORK

Research in usability evaluation is very old and still is an active area of research. Literature provides the evidence that many studies till date have focussed on the various usability evaluation issues and challenges. Studies addressing the contextual issues during usability evaluations can be found in the literature. We reviewed some of the previous work specifically concentrating on the physical and social context in usability evaluations. The related work is classified based on the context they target.

Based on physical context

The importance of physical context in usability evaluations have been researched for a long. Out of the many factors that can effect usability evaluations, physical context is considered to directly influence the behaviour of the people involved in the usability evaluations. The physical context may include the location, the temperature, the time, the light etc.

Tsiaousis & Giaglis [32] examined the effects of environmental distractions on mobile website usability. They proposed a model hypothesizing on the effects of environmental distractions on the usability of mobile sites. They categorized the environmental distractions into auditory, visual and social. A preliminary test on 20 users was conducted to investigate the effect of environmental distractions on mobile website usability. Results confirmed that environmental distractions have direct effect on mobile website usability.

Hummel et al. [33] developed a mobile context-framework based on a small wireless sensor network, to monitor environmental conditions such as light, acceleration, sound, temperature, and humidity during the usability experiments. User experiments have been conducted in a laboratory with seven test persons where the environmental conditions were changed. Under varying environmental conditions the performance of the users on the average was decreased in terms of higher error rates and delays.

Kaikkonen et al. [34] carried out usability testing of mobile consumer application in a laboratory and in a field with a total of 40 test users. Results indicate that field test may be time consuming and may not be a viable option for searching the user interface flaws to improve user interaction. However, they found that field testing is worthwhile when combining usability tests with a field pilot or contextual study where user behaviour is investigated in a natural context.

Razak et al. [23] conducted usability testing with children in both laboratory and field. Drawing applications were tested in their preschool and an educational game was tested in the usability laboratory. The results indicate that field study is more suitable for understanding children experience with technology than it is with testing for usability problems and laboratory study is more suitable for evaluating user interfaces and interaction with the application than it is with understanding children's experience.

Andrzejczak & Liu [35] examined the effect of location on the user's stress level during usability evaluation. User stress levels were assessed by Spielberger's State-Trait Anxiety Inventory; using the paper survey's baseline and experimental stress scores. In addition, user performance data was recorded through task times and subjective user assessments. The data suggested no significant differences exist between participant data in both baseline and experimental anxiety scores. This implies that remote testing as a cost-efficient way to conduct user testing, may be a viable alternative to traditional lab testing without altering the test's effectiveness.

Madathil [36] performed a synchronous remote usability test using a three-dimensional virtual world, and empirically compared it with WebEx, a web-based two-dimensional screen sharing and conferencing tool, and the traditional lab method. The results suggest that virtual lab method is as effective as the traditional lab and WebEx based methods in terms of the time taken by the test participants to complete the tasks and the number of higher severity defects identified. Test participants and facilitators alike experienced lower overall workload in the traditional lab environment than in either of the remote testing environments.

Baillie & Schatz [37] evaluated a multimodal mobile application for a range of mobile devices such as the PDAs and Smart phones. They performed a combination of laboratory and field studies. The users were given a set of four action scenarios to be performed. The results were surprising; only one action scenario was completed in the time frame whereas three out of four action scenarios were completed in lesser time. Error rates were higher in lab than in the field. The reason for such performances by the users could be that the users feel more relaxed in the field.

4.1 Based on the social context

The effect of social context in usability evaluations can be understood by examining some of the previous works in the domain. Work by Benedikte et al. [38] conducted usability testing with 60 children in three setups. Each setup applied think-aloud or constructive interaction method on children in acquainted and non-acquainted pairs. The results show that the pairing of children had impact on how the children verbalized and collaborated in pairs during the testing sessions. The children in pairs had a high level of verbalization, but often they were more talking aloud than actually thinking aloud. The acquainted dyads were significantly more satisfied with their own performance and they did not feel it demanded a lot of effort for them.

Study by Jacobsen et al. [39] examines the evaluator effect in the usability tests. In their study four HCI research evaluators, all familiar with the theory and practice of usability analyzed four video tapes. The evaluators were asked to detect and describe all problems in the interface based on analyzing the four tapes in a preset order, without any time constraints. The results indicate that only 20% of the 93 unique problems were detected by only a single evaluator. Severe problems were detected by more often by all four evaluators (41%) and less often by only one evaluator (22%), however, the evaluator effect remained substantial.

Evaluator effect has also been probed by Hertzum and Jacobsen [40], where they found that different evaluators evaluating the same system with same usability evaluation methods detect substantially different sets of usability problems in the system.

van den Haak & de Jong [41] analyzed the interaction between test monitor and participants in concurrent think aloud (CTA) method and constructive interaction (CI) test. The results indicate

that the presence of the test monitor has most notably affected the CTA participants but has also has its impact on CI participants. They found that a more serious threat to the validity of both the CTA method and CI method was that the participants acknowledged the test monitor as an evaluator of their actions.

5. DISCUSSION

We reviewed many works which took into consideration the importance of the physical as well as the social contexts in usability evaluations. The summary of the relevant work is given in table 2. The entries in the table depict the type of context that was addressed in the work, whether context affected the outcomes of the work and the target participants. Physical context in the table refers to both the laboratory evaluation and field evaluation. Social context refers to the people involved in the test. Both refer to the physical and the social context taken together.

Table 2: Summary of the relevant work

Work	Context used			Does context effects the results			Test participants		Application tested
	Physical	Social	Both	Yes	No	Sometimes	Children	Adults	
[32]	✓	✗	✗	✓	✗	✗	✗	✓	Mobile website
[33]	✓	✗	✗	✓	✗	✗	✗	✓	Mobile
[34]	✓	✗	✗	✓	✗	✗	✗	✓	Mobile commerce application
[23]	✓	✗	✗	✓	✗	✗	✓	✗	Drawing application
[35]	✓	✗	✗	✗	✓	✗	✗	✓	Commercial website
[36]	✓	✗	✗	✗	✗	✓	✗	✓	E-commerce website
[37]	✓	✗	✗	✗	✓	✗	✗	✓	Multimodal mobile application
[38]	✗	✓	✗	✓	✗	✗	✓	✗	Mobile phone
[39]	✗	✓	✗	✓	✗	✗	✗	✓	Video tapes
[41]	✗	✓	✗	✓	✗	✗	✗	✓	Online library catalogues

Physical context is addressed by many studies whereas, the social context in usability evaluations is explored by a fewer studies. Most of the studies show that the context (physical and social) has an impact on the outcomes of the evaluations. The focus on physical context in usability evaluations indicates that this aspect of context is considered important. The choice is between evaluating in an artificial setting such as the laboratory or in a more natural setting through field evaluations. Each has its own strengths and weaknesses.

Jones & Marsden [42] state that the social context in usability evaluation is equally important. The main objective of creating a social context in usability evaluation is to facilitate effective and efficient evaluations. Pardo et al. [43] examined the effect of teacher’s involvement in

usability testing with children. They found that children cannot provide proper feedback on the learning goals they have not experienced before. In such cases the input of the other stakeholders such as the teachers will be beneficial. Some studies also reported that, compared to evaluations involving adult participants, studies involving children are mostly affected by the context. Children show varying behaviour when they are tested in the laboratory environment and when they are tested in the field. They feel relaxed and confident in their own environment. Children also show varying behaviour when they are accompanied by the people they are well acquainted with and with the people they are not acquainted with. Almost 80% of the work that we surveyed shows a clear impact of the context on the outcomes of the usability evaluations. Studies concerning the mobile applications usability testing are the most affected by the context. With mobile applications, better insight to the usability problems can be unveiled in the field than in the lab. None the less, usability evaluations involving adults may also be effected by the context.

While previous research studies in usability evaluations have largely focused on the physical aspects of context in usability evaluations, work on social aspects of context is scarce. Behaviour of the test participants is affected by the context which may in turn affect the results of usability evaluations. Therefore, behaviour and settings are inseparable units as claimed by Barker's theory of behaviour settings. More research taking into account both the social and physical context together is needed to uncover the importance of context in usability evaluations.

6. CONCLUSION

This paper is a review of the role of context in usability evaluations. Essentially the role of the physical and social context in the literature was examined. Roger Barker's theory of behaviour settings can be a strong motivation for investigating context. Behaviour setting theory considers behaviour and the settings are inseparable units .Behaviour setting theory has been widely cited for its potential applications to community psychology. While the theory has broad applicability and a strong empirical base, research on it is limited. Future research will focus on examining the elements of physical and social context in an attempt to understand the influence of context in usability evaluations. However, there is a lack of coherence in understanding the context through psychological perspective. The research on context is scattered and scarce, lacking a unifying overview. Therefore, an understanding of the influence of context and how it impacts the process of usability evaluation is needed. This study has formed the basis for our future work in exploring the role of context in usability evaluations. It has also established the line of investigation that is needed to move forward in developing a usability evaluation framework encompassing context.

REFERENCES

- [1] Nielsen, J. (1993) *Usability Engineering*. San Francisco, Morgan Kaufmann
- [2] Anjoo Navalkar Usability engineering - quality approach (iso 13407). Accessed 10-Dec-2011, <http://www.humanfactors.com/downloads/documents/UsabilityISO.pdf>.
- [3] Bevan N. (1995) "Measuring usability as quality of use", *Journal of Software Quality*, Issue 4, pp115-140
- [4] Nielsen, J. (2003) "Usability 101: Introduction to Usability", Alertbox, Accessed 5-Dec-2011, <http://www.useit.com/alertbox/20030825.html>
- [5] Scholtz, J.(2004) "Usability Evaluation" , in National Institute of Standards and Technology.
- [6] Holzinger, A.(2005) " Usability Engineering Methods for Software Developers", *CACM* Vol. 48, No.1.
- [7] Kantner, L. and Rosenbaum, S.(1997) "Usability studies of WWW sites: Heuristic evaluation vs.

- laboratory testing,” in *Proc. ACM SIGDOC*, pp. 153–160.
- [8] Olmsted-Hawala, E. , Murphy, E., Hawala, S., and Ashenfelter, K. (2010) “Think-aloud protocols: a comparison of three think-aloud protocols for use in testing data-dissemination web sites for usability,” *In Proceedings of the 28th international Conference on Human Factors in Computing Systems*, pp. 2381-2390.
- [9] Schoggen, P. (1989) *Behaviour Settings: A Revision and Extension of Roger G. Barker’s ‘Ecological Psychology’*. Stanford, CA: Stanford University Press.
- [10] Scott, M. M. (2005). A powerful theory and a paradox: Ecological psychologists after Barker. *Environment and Behavior*,
- [11] Barker, R. G. (1968) *Ecological psychology*. Stanford, CA: Stanford University Press.
- [12] Barker, R. G. (1987). Prospecting in environmental psychology: Oskaloosa revisited. In D. Stokols&I. Altman (Eds.), *Handbook of environmental psychology* (Vol. 2).New York: John Wiley.
- [13] Brown, P.J., Bovey, J.D. & Chen, X. (1997) “ Context-Aware Applications: From the Laboratory to the Marketplace”, *IEEE Personal Communications*, 4(5), 58-64. IEEE Press.
- [14] Ryan, N., Pascoe, J. & Morse, D. (1997) “ Enhanced Reality Fieldwork: the Context- Aware Archaeological Assistant”, In Gaffney, V., van Leusen, M. & Exxon, S. (Eds.) *Computer Applications in Archaeology*.
- [15] Hull, R., Neaves, P., Bedford-Roberts, J. (1997) “Towards Situated Computing”, *1st International Symposium on Wearable Computers*, pp. 146-153.
- [16] Schilit, B. & Theimer, M. (1994) “Disseminating Active Map Information to Mobile Hosts”, *IEEE Network*, 8(5), 22-32. IEEE Press.
- [17] Dey, A. K. & Abowd, G. D. (2000) “Towards a Better Understanding of Context and Context-Awareness”, *In CHI2000 Workshop on What, Who, Where, When and How of Context-Awareness*, New York: ACM Press.
- [18] Rosenbaum, S. (1989) “Usability evaluations versus usability testing: When and why?,” *IEEE Trans.Prof. Commun.*, vol. 32, no. 4, pp. 210–216.
- [19] Bevan, N. & Macleod, M. (1994) “Usability Measurement in Context”, *Behaviour and Information Technology (BIT)*, 13(1-2), pp.132-145.
- [20] Nivala, A.-M. & Sarjakoski, L. T. (2003) “Need for Context-Aware Topographic Maps in Mobile Devices”, *Proceedings of the 9th Scandinavian Research Conference on Geographical Information Science (ScanGIS’03)*, pp.15-29.
- [21] Maguire, M. (2001) “Context of Use within Usability Activities”, *International journal of Human-Computer Studies*, 55(4), pp.453-483, Taylor & Francis Group.
- [22] Nielsen, C.M., et al. “It's Worth the Hassle! The Added Value of Evaluating the Usability of Mobile Systems in the Field”, *Nordi CHI’06*. ACM.
- [23] Razak, F.H.A., Hafit, H., Sedi, N., Zubaidi, N.A.,Haron, H. (2010) "Usability testing with children: Laboratory vs field studies”, *User Science and Engineering (i-USEr), 2010 International Conference on* , vol., no., pp.104-109.
- [24] Johnson, P. (1998) “ Usability and Mobility; Interactions on the Move”, *First workshop on human - computer interaction with mobile devices. GIST Technical Report G98-1.* (<http://www.dcs.gla.ac.uk/~johnson/papers/mobile/HCIMD1.html>)
- [25] Kjeldskov, J. and Graham, C. (2003) “A review of Mobile HCI Research Methods”, *In Proceedings of Mobile HCI 2003*, Springer-Verlag, LNCS 2795, 2003, pp. 317-335.
- [26] Park K.S. and Lim C.H. (1999), “ A structured methodology for comparative evaluation of user interface designs using usability criteria and measures”, *International Journal of Industrial Ergonomics* 23, pp. 379 - 389.
- [27] Markopoulos, P., Read, J. C., MacFarlane, S. & Höysniemi, Johanna (2008) “Evaluating Children’s Interactive Products: Principles and Practices for Interaction Designers”, Burlington, MA: Morgan Kaufmann.

- [28] Oh, K., & Kim, M.(2010) "Social Attributes of Robotic Products: Observations of Child-Robot Interactions in a School Environment", *International Journal of Design*, vol 4 no.1 pp.45-55.
- [29] Xu, D., Mazzone, E. and MacFarlane, S.(2006) "In Search for Evaluation Methods for Children's Tangible Technology", in *Proceeding of IDC Conference 2006*, Tampere, Finland.
- [30] Stoica, A., Fiotakis, G., Simarro, J., Muñoz, H., Avouris, N., Dimitriadis, Y. (2005) "Usability evaluation of handheld devices: A case study for a museum application", *In Proceedings PCI 2005* November 2005.
- [31] Jane, J.J.(2009) " Social Context in Usability Evaluations: Concepts , Processes and Products ", PhD thesis. (http://vbn.aau.dk/files/44026023/Jul_Jensen_PhD_thesis.pdf)
- [32] Tsiaousis, A.S., Giaglis, G.M.(2008) "Evaluating the Effects of the Environmental Context-of-Use on Mobile Website Usability," *Mobile Business, 2008. ICMB '08. 7th International Conference on* , vol., no., pp.314-322.
- [33] Hummel, K. A., Hess, A., and Grill, T. (2008) " Environmental context sensing for usability evaluation in mobile hci by means of small wireless sensor networks", *In MoMM '08: Proceedings of the 6th International Conference on Advances in Mobile Computing and Multimedia*, pp. 302–306.
- [34] Kaikkonen, A., Kallio, T., Kekäläinen, A., Kankainen, A., Cankar, A. (2005) "Usability Testing of Mobile Applications: A Comparison between Laboratory and Field Testing," *Journal of Usability studies* (1:1) 2005, pp 4-16.
- [35] Andrzejczak, C., Liu, D. (2010) "The effect of testing location on usability testing performance, participant stress levels, and subjective testing experience", *Journal of Systems and Software* Volume 83, Issue 7.
- [36] Madathil, K.C. (2011) "Synchronous remote usability testing: a new approach facilitated by virtual worlds", *Proceedings of the 2011 annual conference on Human factors in computing systems CHI'11*.
- [37] Baillie, L., Schatz, R. (2005) "Exploring Multimodality in the Laboratory and in the Field", *Proc. 7th Intern. Conference on Multimodal Interfaces*, pp. 100-107.
- [38] Als, B. S., Jensen, J. J., Skov, M. B.(2005) " Exploring Verbalization and Collaboration of Constructive Interaction with Children", *Proceedings of the IFIP 9th International Conference on Human-Computer Interaction - INTERACT 2005*, pp. 443-456.
- [39] Jacobsen, N., Hertzum, M., and John. B. (1998) "The evaluator effect in usability studies: Problem detection and severity judgments," in *Proc. Human Factors and Ergonomics Society 42nd Annual Meeting*, pp. 1336–1340.
- [40] Hertzum, M., and Jacobsen, N.E. (2001) "The Evaluator Effect: A Chilling Fact about Usability Evaluation Methods", *International Journal of Human-Computer Interaction*, 13.4, pp. 421-443.
- [41] van den Haak, M.J.; de Jong, M.D.T. (2005) "Analyzing the interaction between facilitator and participants in two variants of the think-aloud method," *Professional Communication Conference, 2005. IPCC 2005. Proceedings International* , vol., no., pp. 323- 327.
- [42] Jones, M. & Marsden, G. (2006) "Mobile Interaction Design", West Sussex, England: John Wiley and Sons.
- [43] Pardo, S.; Vetere, F.; Howard, S. (2006) Teachers' involvement in usability testing with children. In: *IDC 2006*