Usability Evaluation in Software Development Practice

Marta Kristin Lárusdóttir

HAL Id: hal-01596971
https://inria.hal.science/hal-01596971

Submitted on 28 Sep 2017

Usability Evaluation in Software Development Practice

PhD Student:

Marta Kristin Larusdottir
Reykjavik University, Menntavegur 1, 109 Reykjavik
The Royal Institute of Technology, Lindstedtsvägen 5, 10044 Stockholm
marta@ru.is

Supervisor:

Dr. Jan Gulliksen,
The Royal Institute of Technology, Lindstedtsvägen 5, 10044 Stockholm
gulliksen@kth.se

Research area: Usability, evaluation, agile development, Scrum, practice.

Brief description of the topic: Usability evaluation is explored by comparing the effectiveness of using different evaluation methods and by studying how usability evaluation is conducted by practitioners.

1 Description of the Background to the Research Topic

Usability evaluation is one of the core usability activities, defined in the ISO 9241:11 standard and in the ISO 13407 standard [5, 6]. During usability evaluation the evaluator actively defines the status of the usability of the system being evaluated. He or she often uses methods or techniques to focus the activities. Usability evaluation has proven to be a difficult task for the evaluators [9] so understanding how different the results are when various usability evaluation methods are used, understanding the context which evaluators work in and the hindrances they are experiencing is important.

Researchers have been interested in estimating the effectiveness of using various usability evaluation methods since early 90's see for example [2, 7, 8] to be able to advice software developers which method to use. The effectiveness of using a particular evaluation method has been measured by comparing the number of usability problems found by using various methods.

Comparing usability evaluation methods by counting usability problems that are found by each method was criticized by Wixon [15], were he states that research results on evaluation methods fails the practitioner, meaning people working actively

on software development. Wixon [15] states that: "the development of real products is the only context sufficiently rich to produce the kind of nuanced examples that are needed to develop a differentiated and contextualized understanding of methods and techniques needed by practitioners". Some researchers have responded to his claim and have done studies that have the goal of supporting practitioners better, for example Uldall-Espersen and collages [14] where they study how useful usability evaluation results are to the practitioners.

Agile processes for software engineering evolved to address perceived limitations of more established, plan-driven approaches for software development and have shown their worth in quickly developing reliable software [1]. The most widely used are XP and Scrum [11]. In Scrum the projects are split up in two to four weeks long iterations called sprints, each ending up with a potential shippable product. In Scrum self organizing and well compounded teams are heavily emphasized, typically with 6 – 8 interdisciplinary team members [12].

One of the benefits of using agile development processes was claimed to be that customers needs are taken more into account than when developing software using more traditional processes [12]. Traditional Scrum has been criticized for not involving real users in their software process and for not adequately address their usability needs, for example in [13]. Evaluation of the success of deploying user involvement methods in Scrum has so far only been anecdotal [11].

2   Description of the Research Topic and my Contribution

My first research question is: *Does the use of different usability evaluation methods give different results when evaluating the same software system?*

Through this question I have explored if the results of doing usability evaluation with users using the think aloud method and the results of using the inspection methods, heuristic evaluation and cognitive walkthrough are different. I have done two experimental studies on this subject, where the quantity of usability problems found by evaluators using these methods is compared [3, 4]. Furthermore, the quantity of serious problems found was studied. The reason for comparing these methods even though these could be used in different purposes is to understand how different the results are and what the strengths and weaknesses are of using these methods. That understanding will help practitioners to choose an evaluation method when evaluating a system they are developing and academics could use the results to be able to suggest new methods or techniques for evaluation [10].

My results from the first study show that the think-aloud method is the most effective method when comparing the results of using the three methods. Similar quantity of problems were found using the think-aloud method and using the heuristic evaluation, but a large extent of the problems found by using heuristic evaluation were false positive, meaning that these were not found during the think-aloud evaluation. In the second study one third of the problems found by using the think-aloud method were found by using the heuristic evaluation and again a fair amount of problems were false problems.

Through my second question: *How do practitioners integrate usability evaluation in the Scrum development process?* I explored how usability evaluation is conducted by software developers using the Scrum process in Iceland. By using a survey I studied how often developers evaluate usability, how that compares to other activities like testing and what the hindrances of doing usability evaluation are. The main findings show that unit, functional, system and acceptance testing are done to a wide extent [9]. Usability, security, performance, alpha and beta testing are much less emphasized. Interviews were conducted to exemplify how practitioners conduct usability testing and what they describe as the difference of usability and acceptance testing. Many of the respondents said that usability testing can only be done once or twice a year. The respondents did not have time for more frequent evaluation. Some examples from the interviews show that practitioners are willing to do formal usability testing on extensive parts of the system, but because the iterations in Scrum are short and the changes to the system in each iteration are small, formal usability testing does not fit into the project work.

Furthermore through my third research question: *What are the challenges for practitioners to integrate usability activities in Scrum development projects?* I have studied how other usability activities are integrated into the Scrum process to understand if the challenges of integrating usability evaluation are similar to the challenges integrating other usability activities. To do this we have conducted an interview study with 21 informants from Swedish software companies that use the Scrum process. The analysis of this study is still in progress, but the first results show that formal usability evaluation is not much emphasized. Instead more informal ways of evaluation are used, like observing users, getting feedback on the usability from users through group meeting, blogs or chats. Furthermore informal expert evaluations like peer reviewing are conducted by a number of the informants and evaluation with user representatives. The main obstacle for evaluating the usability formally with users more frequently than once a year is that the team does not have time. Non-functional requirements like usability are rarely stated or written. Some informants mention that it is hard to know when the usability of the system is good enough; it is always possible to extend the level of the usability. Furthermore the practitioners have different opinions on who is responsible for that the software system is usable, mentioning the whole team as being responsible, the project leader, the usability expert or actually no one as possible options.

To summarize, my research topic is usability evaluation, which is explored by comparing the effectiveness of using different evaluation methods and by studying how usability evaluation is conducted by practitioners using the software development process Scrum. This research topic is important, because even though usability evaluation methods are widely known in the research community of HCI, the methods are not widely used by the practitioners using the Scrum development process. Understanding how different the results of using various evaluation methods and how practitioners are evaluating their systems is of great importance for the HCI research community to be able to suggest new methods or ways to evaluate the usability of software systems that fit the practitioners.

3 References

1. Beck, K. Extreme programming explained: Embrace change, 2nd edition. Reading, MA: Addison Wesley (2004).
2. Desurvire, H. W., Kondziela, J. M., Atwood, M. E., What is gained and lost when using evaluation methods other than empirical testing. In *Proc. People and Computers VII*, Cambridge University Press (1993), 89-102.
3. Frøkjær, E., Larusdottir, M. K. Prediction of Usability : Comparing Method Combinations. In Proc. of Idea Group publishing conf., Hershey, (1999), p. 248 – 257.
4. Hvannberg, E. Þ., Law, E. L., Lárusdóttir, M. K. Heuristic evaluation: Comparing ways of finding and reporting usability problems. *Interacting with Computers*, (2007); 19 (2): 225 - 240.
5. International Organization for Standardization, 1998: ISO 9241-11 Ergonomic requirements for office work with visual display terminals (VDTs) — Part 11: Guidance on usability, Geneva, Switzerland.
6. International Organization for Standardization, 1999: ISO 13407 Human-centred design processes for interactive systems, Geneva, Switzerland.
7. Jeffries, R., Miller, J. R., Wharton, C., Uyeda, K. User interface evaluation in the real world: A comparison of four techniques. In *Proc. CHI 1991*, ACM Press (1991), 119-124.
8. Karat, C., Campbell, R. Fiegel, T. Comparison of empirical testing and walkthrough methods in user interface evaluation. In *Proc. CHI 1992*, ACM Press (1992), 397-404.
9. Larusdottir, M. K. Bjarnadottir, E. R., Gulliksen, J: The Focus of Usability Testing in the Software Industry, Proceedings of the World Computer Congress, Brisbane, Australia, (2010).
10. Larusdottir, M. K. Listen to your Users – The Effect of Usability Evaluation on the Software Development Process, Uppsala University, (2009).
11. Salah, D., Petrie, H., Paige, R. F., Towards a Framework for Integrating User Centered Design and Agile Software Development Processes, *In Proc. Irish CHI, (2009).*
12. Schwaber, K. (1995). Scrum development process. *OOPSLA'95 Worshop on Business Object Design and Implementation.*
13. Singh, M. (2008). U-SCRUM: An Agile Methodology for Promoting Usability. *AGILE '08. Conference*, (pp. 555-560).
14. Uldall-Espersen, T., Frøkjær, E., Hornbæk, K. Tracing impact in a usability improvement process. *Interacting with Computers, 20,* Elsevier B. V. (2008) 48 – 63.
15. Wixon, D. Evaluating usability methods: why the current literature fails the practitioner. *Interactions, 10, 4 (2003), 28 – 34.*