Mednarodno inovativno poslovanje
Journal of Innovative Business and Management

1.01 Original scientific article = Izvirni znanstveni članek

Assessment of a Method for Usability Testing by Determining Usability of the Online Learning Platform

Darko Števančec* | **Iris Fink Grubačević****

Abstract*: Today many products are designed to be reasonably easy to use so, in addition to usefulness, product usability is of great importance. Usability means simply using or simply learning to use a tool or device. In software engineering, usability is the degree to which certain users can effectively use software to achieve their goals in the context of use. Usability assessment involves measurement methods, such as analyzing consumer needs and examining perceptions of product efficiency and product attractiveness. In human-computer interaction, studying usability means the analysis of attractiveness and clarity of an application which is designed to interact with a computer program or web site. Usability takes into account both user satisfaction as well as the usability and quality of the component under study, while striving to improve the user experience through iterative product design. The basic problem of the research project was the pilot assessment of the suitability of our own method of testing the usability of the Sakai online learning platform in integration with the BigBlueButton conference call program. The article presents the most important results of assessing our own method of testing the usability of the learning platform on groups of students and lecturers participating in the e-learning process and project (CPK) testers. The testing was performed on the basis of the evaluation of the test subjects' introspective insight into their own expression of emotional reactions by means of a questionnaire and analysis of mimics, gestures and proxemics from the videos of the test subjects by the testers. The statistical significance of the differences between the groups was performed by Kullback 2Î test of the independence hypothesis.*

Keywords: method; usability; e-learning

JEL classification: C19

*Assist. Prof. Dr., DOBA Business School, Prešernova ulica 1, 2000 Maribor, Slovenia; darko.stevancec@net.doba.si

** Sen. Lect., MA, Faculty of Industrial Engineering Novo mesto, Šegova ulica 112, 8000 Novo mesto, Slovenia; iris.finkgrubacevic@fini-unm.si

How to cite this paper = Kako citirati ta članek: Števančec, D. and Fink Grubačević, I. (2019). Assessment of a method for usability testing by determining usability of the online learning platform. *Mednarodno inovativno poslovanje = Journal of Innovative Business and Management*, 11(3), 35-41. DOI: 10.32015/JIBM/2019-11-3-4

Mednarodno inovativno poslovanje = Journal of Innovative Business and Management
ISSN: 1855-6175

Preverjanje ustreznosti spletne učne platforme z lastno metodo testiranja uporabljivosti

Povzetek: *Danes so številni izdelki zasnovani tako, da jih je predvidoma enostavno uporabljati, zato je poleg uporabnosti izjemnega pomena tudi uporabljivost izdelka. Uporabljivost pomeni preprosto uporabo ali preprosto učenje uporabe orodja ali naprave. V programskem inženiringu je uporabljivost stopnja, do katere lahko določeni uporabniki učinkovito uporabljajo programsko opremo za doseganje ciljev v kontekstu uporabe. Ugotavljanje uporabljivosti*

vključuje metode merjenja, kot so analiza potreb potrošnika in preučevanje percepcije učinkovitost uporabe ter privlačnosti izdelka. V interakciji človeka z računalniki preučevanje uporabljivosti pomeni analizo privlačnosti in jasnosti aplikacije, s katero je oblikovana interakcija z računalniškim programom ali spletno stranjo. Uporabljivost upošteva tako zadovoljstvo uporabnikov kot tudi uporabnost in kakovost proučevane komponente ter si hkrati prizadeva izboljšati uporabniško izkušnjo z iterativnim oblikovanjem izdelka. Osnovni problem raziskovalnega projekta je bil pilotsko preverjanje ustreznosti lastne metode testiranja uporabljivosti spletne učne platforme Sakai v integraciji s programom za konferenčne klice BigBlueButton. V članku so predstavljeni najpomembnejši rezultati preverjanja lastne metode testiranja uporabljivosti učne platforme na skupinah študentov in predavateljev, ki sodelujejo v procesu e-študija ter projektnimi (PKP) testatorji. Preverjanje smo izvedli na podlagi evalvacije introspektivnega vpogleda testirancev v lastno izražanje emocionalnih reakcij s pomočjo vprašalnika in analize mimike, gestike in proksemike iz video posnetkov preizkušancev s strani testatorjev. Ugotavljanje statistične pomembnosti razlik med skupinami smo izvedli s Kullbackovim 2Î preizkusom hipoteze neodvisnosti.

Ključne besede: *metoda; uporabljivost; e-študij*

1 Introduction

Usability is one of the most important features of a product which also defines its value on the market. People have been developing and marketing products for a long time so the International Organization for Standardization (ISO) has developed the 9241-11 standard. Official ISO 9241-11 standard states that usability is the ability to use a particular user's product for a specific purpose which is met with efficiency, performance, and satisfaction in a particular context of use (ISO, 2018).

Usability applies to all aspects of use, including the ability to teach new users the effectiveness, productivity and satisfaction in learning how to use the new system. It can be considered as one of the quality parameters that shows how easy a particular product is to use. This word also refers to the methods by which we fasten or improve the ease of use during the product design process itself. It can be defined with the following components: learnability - how fast the user can learn how to perform basic tasks when they first learn about the design; efficiency - when the user learns design, how fast they can perform tasks; recall - when the user has not used the program for some time, how long it will take to regain knowledge; mistakes - how many mistakes the user makes, how big those mistakes are and how much effort he has to make to solve them; satisfaction - what is the satisfaction level while using the design (Nielsen, 2019).

Barnum (2010) highlights three critical elements in defining usability: specific users - not any user but specific users for whom the product was designed, specific goals - specific user goals are aligned with product goals, and specific use framework - the product must be designed in such a way to work in an environment where users will use it.

1.1 Product usability

According to external users, the usability of the product is extremely important as it is one of the key factors for survival in the market. If the product is unusable, too slow or too complex, we run the risk that users will not use it. Unless it is clearly stated what features the product offers, users may lose interest. Likewise, with a lack of ease in reviews and answers to questions the interest of users fades. When using a product in business, productivity may fall due to unadopted design and reduced usability (Nielsen, 2019).

Usability is a quality that many products have. It is difficult to define what makes a product usable, to put it bluntly, a product is usable when the user is not frustrated when using it (Rubin & Chisnell, 2008). The definition of usability, according to Rubin & Chisnell (2008), is that a product is really usable when the user can do what he wants with the product or expects the product to use without hindrance, hesitation or question. Usability is problematic when it is lacking or absent from the product.

Quesenbery (2001) identifies four essential characteristics of usability: efficiency, performance, inevitability/friendliness, and ease of memorization. Efficiency is the perfection and precision with which users reach

certain goals. It is judged whether the goals of the user have been successfully fulfilled and whether all the work has been completed. The help built into the system, which guides the user and helps him to fix a problem, also has a major impact on the efficiency of the user's experience. Performance can be described as the speed and accuracy at which users can complete the tasks for which they use the product. ISO 9241 defines, among other things, efficiency as the total resources consumed in a task. Navigation design elements, such as keyboard shortcuts, menus, links, and other buttons, affect performance because, if well designed, with clear actions, less time and effort is required for the user to select navigation and actions. Inevitability (friendliness) is achieved through pleasant experience and satisfactory use. The visual presentation style, number, features and types of graphic images or colors and the use of all multimedia elements are part of the user's immediate response. It is important that the design meets the expectations and needs of the people using the product. Creating an »introduction« to a product that explains the features and tools, with the right language and the right visual aids, can help improve ease of memorization.

Visual factors influence the overall user experience. As Soegaard (2019) points out, special attention should be paid to the size and color of the font which is easy to read, the color layout to develop aesthetic appeal, while providing readability, navigation for users to get the most out of the site point A (point of entry) to point B (where they want to be) as quickly and easily as possible, content where design elements must be tailored to the way information is presented, to the user experience, to the goal the product pursues and to titles that are meaningfully organized.

There are other important quality parameters. One of the key parameters is the utility that relates to the functionality of the design. Usability and usefulness are both very important as they together determine whether a design is useful. It is also not good that the system can do everything you can imagine, but it turns out that the complexity of it makes it almost impossible to perform. The methods we use to study the usefulness of design can also be used to analyze and improve usability (Nielsen, 2019).

The concept of utility is defined provided that the program/system represents important tools for completing the user's work. Utility answers the question of whether the system supports the user's need for information and his work. The user searches the system for appropriate resources or information for his work tasks. This information, which the tasks require, needs formulation, expression, inquiry, relevance assessment tied all together and executed in a repetitive manner. During these stages of information retrieval common features are expressed here as requirements or criteria, depending on the level of the query. The ITF (Interaction Triptych Framework) summarizes these features in relevance, reliability, information level, format and time coverage. The appearance of the system also plays a significant role, as aesthetic appearance and layout play a key role in the ultimate level of satisfaction (Tsakonas & Papatheodorou, 2008).

Ease of use in a given context is determined by the attributes of the product and is measured by efficiency and user satisfaction. The context consists of the user, the task, and the physical and social environment. Product attributes that contribute to usability include the style and properties of the product, the structure of the user-product dialogue, the nature of the functionality, and any other important attributes such as efficiency and reliability. Attitude and performance criteria provide criteria that determine if product design is successful in achieving usability. In the future analytical techniques will be able to predict the relationship and performance of these properties. The system to be used in the real world must be acceptable to the individual user - the author refers to this as »utility« (Bösser, 1991).

Queiroz (2017) states that usability is a specific aspect of software development and design, is considered one of the features of a standalone program and can be defined as »the extent to which a particular user can use a product to achieve certain goals with efficiency, performance and satisfaction in a particular context«. However, without proper usability of the program, we cannot market it, customers cannot even use it within the target group.

1.2 Methods of usability measurement and usability testing

Testing mainly focuses on the user's ease of use of the application, flexibility in management and the ability of the system to achieve its goals. It is also called UX - User ExperienceTesting. Usability testing is a fundamental method of determining the usability of a product. The basic idea is that under controlled circumstances users perform tasks that are predetermined while closely monitoring their behavior, functioning, reactions, and recording their performance (Nielsen, 1993).

Usability testing is a method by which product users are required to perform specific tasks in order to measure the ease and efficiency of product use, time to complete a task, user perception and experience. Usability testing can be done formally with video cameras or informally with a paper model of the application. Based on the usability tests changes which build on the improvement of user's experience are made. Whether the test is formal or informal test participants are encouraged to think aloud and express their opinions. Usability testing is best done in conjunction with the design (purpose) of the application and the focus on the user (Rouse, 2005).

Research (Virzi, 1992; Nielsen, 1994; Nieslen and Landauer, 1993) shows that five users are sufficient to detect 80% of usability problems. The actual number of users required depends on the complexity of the application, the objectives of the test and the expected end-use of the product.

There are many usability analysis methods, the most useful and basic is user testing: we need to get relative customers, we assign them tasks that they need to repeat several times and we mark the findings of what users are doing and where they are having problems with our product. It is important to let the users talk. During the test it is also important to let users solve the problem on their own and not help them with it, as this can destroy the test. In order to find out what problems our product has, it is enough to have five different people test it, so we do not need to invest a lot of money for extensive and expensive analyzes. The more product versions we test, the better for us (Nielsen, 2019).

If the test is performed using a method in the presence of the observer, care should be taken to ensure that the context of use matches the intended context in the real world (user types, tasks and environment) as closely as possible. Controlled measurement always obscures certain aspects of the real work environment, which should be taken into account when analyzing data (Bösser, 1991).

When we talk about the usability test, we are talking about an activity that focuses on observing users using the product and performing real-world tasks that are meaningful for that product. Using this definition, usability tests can be divided into formative testing - while a product is in development, our goal is to diagnose and identify problems; it is based on small, repeated studies during product development. Summative Testing - upon completion, our goal is to determine baseline measurements or confirmation that the product meets the requirements; it usually requires a larger number of users for statistical validity (Barnum, 2010).

The primary objective of the usability test is to improve the usability of the product being tested. The secondary goal is to improve the process by which products are designed and developed in order to avoid the same problem with other products. These characteristics help us distinguish the usability test from the research (Dumas and Redish, 1999).

The only way to have a quality user experience is to start testing with users at an early stage of design and continue to test every step of the way. For most businesses, it is good to have the tests done in a meeting room or office behind a closed door, protected from external interference. It is important to get the right users and sit with them while using the test format. Recording feedback is vital.

It is necessary to decide on one of the above methods. The goal pursued by product enhancements is that it must be usable in the real world and be acceptable to every single user, who must judge that the benefits of using this product are greater than any other means of achieving the task/goal. Acceptability will depend on the context of use and the characteristics of the user, which may be influenced by factors such as cost, convenience, availability, pre-training, affection for computers or other organizational constraints.

2 Method

Assessment of the usability of the online learning platform with our own usability testing method was carried out as a part of the project “Creative path to knowledge”. Assessment of the method suitability was carried out through the following successive project phases:

– One-month training of 7 CPK testers in the fields of non-verbal communication and recognition of emotional reactions. The project phase included the study of scientific and professional literature as well as individual training in the acquisition of emotional reaction recognition skills.

- Construction of a simple measuring instrument - a questionnaire with a complex set of expressions of 37 emotions/reactions: astonishment, joy, surprise, thoughtfulness, rage, apathy, euphoria, boredom, doubt, horror, anger, calmness, tension, disgust, insecurity, despair, distrust, enchantment, dissatisfaction, relief, impatience, resistance, pride, curiosity, hatred, excitement, relaxation, disappointment, fear, stubbornness, pleasure, agitation, interest, sadness, distress, bewilderment and alertness.
- Test of the knowledge, abilities and skills of CPK testers in recognizing emotional reactions by analyzing prepared video inserts of persons in different emotional states/reactions.
- Video recording of 11 test subjects (students and professors) testing the Sakai online learning platform, and introspective assessment of the expression of individual test subject's emotions when testing a product by using a survey questionnaire.
- Analysis of test subjects video recordings by 7 CPK testers and assessment of the expression of individual emotional reactions by means of a questionnaire (students and professors).

To determine the relationship between the two variables (status and emotion), Kullback 2Î-values were calculated because the theoretical frequencies in the individual cells of the contingent table were less than 5 and it was not possible to use the Chi-squared test. Differences in the perception of emotional responses were identified between two independent samples: 7 CPK testers and 11 test subjects (professors and students participating in e-learning). A potential limitation of the research is the question of the relevance of the introspective insight of the test subjects, which will need to be verified with additional statistical analyzes of the consistency of the respondents' assessments. Data processing was performed with the IBM SPSS 19 statistical package.

3 Results

In the chapter we will present the most important results of a pilot study, namely Kullback 2Î-tests of the correlation of two variables (status and emotion), whereby testing the independence hypothesis we found statistical differences, thus refuting the independence hypothesis.

Table 1: Kullback 2Î-test (status, emotion)

	Value	df	Sig. (2-tailed)
Kullback test - thoughtfulness	22.074	3	.000
Number of recognitions	17		
Kullback test - doubt	14.045	4	.007
Number of recognitions	13		
Kullback test - calmness	21.170	5	.001
Number of recognitions	16		
Kullback test - relief	4.499	1	.034
Number of recognitions	4		
Kullback test - curiosity	17.006	2	.000
Number of recognitions	17		
Kullback test - interest	25.864	4	.000
Number of recognitions	19		
Kullback test – alertness	17.397	3	.001
Number of recognitions	15		

* α=0,05

The most recognized emotional reactions present are in the following order: interest, thoughtfulness, curiosity, calmness, alertness, doubt and relief. The result of the Kullback 2Î test of the independence hypothesis is statistically

significant at the α = 0.05 level (p = 0.001, p <0.05) in the emotional reaction of interest. This emotion was most commonly recognized by both the CPK group of testers and test subjects (professors and students). While analyzing the test videos, this emotion was detected in all 11 test subjects by CPK testers and 8 test subjects also recognized interest in themselves during the test. Altogether, an emotion of interest was observed in nineteen cases. The value of the Kullback 2Î test of the independence hypothesis in emotional reaction of thoughtfulness is statistically significant (p = 0.000). This emotion was observed by the testers in all 11 test subjects while analyzing the videos, and 6 test subjects observed thoughtfulness in themselves during the testing. The emotion of thoughtfulness was thus recognized 17 times. Equally common was the perception of curiosity, CPK testers have observed it in ten subjects and introspectively it was recognized in 7 test subjects.

The emotional response of calmness was recognized 16 times in total, of which CPK testers have recognized it ten times and 6 test subjects have reported to have been calm when interacting with the tested platform. The Kullback 2Î test was statistically significant (p = 0.001). Emotion/reaction of alertness has a similar high visibility (Kullback 2Î-test p = 0.001), it was recognized in 15 cases in total, with CPK testers recognizing it in all 11 test subjects and 4 test subjects rated themselves as being alert during the platform testing. Relatively high emotion recognition occurs with emotional reaction of doubt. Doubt was recognized 13 times in total, self-recognized by three test subjects, and recognized in 10 subjects by CPK testers (p = 0.007).

Relief was the least recognized emotional reaction, with only four cases in total. It was self-recognized by just one test subject and three CPK testers recognized the relief based on video analysis. Kullback 2Î test of the independence hypothesis is statistically significant at the α = 0.05 level (p = 0.034, p <0.05).

4 Discussion and conclusion

The results of a pilot study of assessment of our own method for analyzing the emotional reactions of test subjects in contact with a tested product, based on human analysis of videos of product usability testing, are promising. The test results of the method indicate satisfactory reliability of the method with the limitations so far, such as the relatively small sample of testers and test subjects and the relatively short training of the testers. The method has some comparative advantages over software video analysis of non-verbal reactions, such as a lower cost of analysis and, above all, a higher range of the possible number of emotional reactions analyzed. On the other hand, our own method also has drawbacks, such as the reliability of the results depends on the experience of the testers and the longer latency of obtaining the results (not in real-time as in software analyzes).

We conclude that by upgrading the method of analysis of satisfaction, usefulness and quality, it is possible to obtain reliable results of assessing the usability of a specific product. However, the development of the method also opens up a wide field of application in other areas beyond product usability testing, such as political marketing, business negotiation, criminal forensics and other areas.

References

1. Barnum, C. M. (2010). *Usability testing essentials: ready, set... test!* Burlington, MA: Elsevier.
2. Bösser, T. (1991). Techniques for predicting usability. In: H.J. Bullinger (Ed.), *Proceedings of the 4th International Conference on Human Computer Interaction.* Stuttgart: Elsevier.
3. Dumas, J. S., Redish, J. (1999). *A practical guide to usability testing*. Exeter and Portland, OR: Intellect books.
4. ISO (2018). ISO 9241-11:2018(en). Ergonomics of human-system interaction. Available from: https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en [accessed: March 13, 2019].
5. Nielsen, J. and Landauer, T. K. (1993). A mathematical model of the finding of usability problems. In: *CHI '93 Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, pp. 206-213. Available from: https://dl.acm.org/citation.cfm?id=169166 [accessed: March 12, 2019].
6. Nielsen, J. (1994). *Usability Engineering*. Boston [etc.]: AP Professional.
7. Nielsen, J. (2012). Usability 101: Introduction to Usability. Available from: https://www.nngroup.com/articles/usability-101-introduction-to-usability/ [accessed: March 13, 2019].

8. Queiroz, F., Silva, R., Miller, J., Brockhauser, S., and Fangohr, H. (2017). *Track 1 Paper: Good Usability Practices in Scientific Software Development*. 10.6084/m9.figshare.5331814.
9. Quesenbery, W. (2001). What does usability mean: Looking beyond 'ease of use'. WQusability [online]. Available from: https://www.wqusability.com/articles/more-than-ease-of-use.html [accessed: March 12, 2019].
10. Rouse, M. (2005). Usability. *WhatIs.com* [online]. Available from: https://whatis.techtarget.com/definition/usability [accessed: March 12, 2019].
11. Rubin, J., and Chisnell, D. (2008). *Handbook of usability testing: how to plan, design and conduct effective tests*. Indianapolis, IN: John Wiley & Sons.
12. Soegaard, M. (2019) Usability: A part of the User Experience. *Interaction Design Foundation* [online]. Available from: https://www.interactiondesign.org/literature/article/usability-a-part-of-the-user-experience [accessed: March 12, 2019].
13. Tsakonas, G., & Papatheodorou, C. (2008). Exploring usefulness and usability in the evaluation of open access digital libraries. *Inf. Process. Manage. 44*, 1234-1250.
14. Virzi, R. A. (1992). Refining the Test Phase of Usability Evaluation: How Many Subjects Is Enough? *Human Factors: The Journal of the Human Factors and Ergonomics Society 34(4), 457-468.*