SUMI Evaluation of the EducaNext Educational Portal

TANJA ARH[1], MATJAŽ DEBEVC[2], MATIJA PIPAN[1], BORKA JERMAN BLAŽIČ[1]

1 Laboratory for Open Systems and Networks
Jožef Stefan Institute
Jamova 39, 1000 Ljubljana
SLOVENIA
http://www.e5.ijs.si

2 Faculty of Electrical Engineering and Computer Science
University of Maribor
Smetanova ulica 17, 2000 Maribor
SLOVENIA
http://www.feri.uni-mb.si

Abstract: - The new challenge for designers and Human Computer Interaction researchers is to develop software tools and applications for effective e-learning. Usability is one of the most important aspects of software applications. However, in practice, not much attention is given to this issue during testing. Evaluators often do not have the knowledge, instruments and/or time available to handle usability. The paper presents the results of empirical study of usability evaluation which was based on SUMI (Software Usability Measurement Inventory) questionnaires. The software application tested was a multilingual educational portal EducaNext.

Key-Words: - human computer interaction, EducaNext educational portal, distance learning, software usability measurement inventory (SUMI)

1 Introduction

The usability testing is of key importance in the human-computer interaction. It is one of the basic elements used to verify the user interface quality [7]. There are many definitions of usability. All of the definitions, including ISO 9241/11[4], consider multiple factors, such as ease of learning, ease of use, effectiveness of the system, user satisfaction; the definitions connect those factors to the impact on humans. There are many evaluation methods [8] used in usability evaluations. The method selection often depends on what is being evaluated, the software and hardware used, users that are tested and the research budget.

In our case, we used the Software Usability Measurement Inventory (SUMI) method [1], which was developed in the project 'Metrics for Usability Standards in Computing' (MUSiC, CEC ESPRIT project number 5429) by the Human Factors Research Group (HFRG), University College, Cork. This generic usability tool is comprised of a validated 50-item paper-based questionnaire in which respondents score each item on a three-point scale (i.e., agree, undecided, disagree). The questionnaire is designed to measure the affect, efficiency, learnability, helpfulness and control. During its development, the questionnaire was standardized as a measurement tool for some of the user-orientated requirements expressed in the European Directive on Minimum Health and Safety Requirements for Work with Display Screen Equipment (90/270/EEC). SUMI is also mentioned in the ISO 9241 standard as a recognized method of testing user satisfaction.

Users normally need about ten minutes to complete the inventory. In a software development environment if the users have no previous experience of the software, additional time is needed for introduction, training, and carrying out a set of benchmark taks with the software. How long this takes depends on the complexity of the software being evaluated and may be from 20 minutes to more than an hour.

Analyzes and testing have been realized within the Phare FOCUS-SIAT project – e-learning and training in the field of cross border cooperation aimed at the strengthening of the management proficiency levels in small and medium size enterprises and thus contributing to the heightening of the educational attainment of employees at the bordering regions between Slovenia and Austria.

In the next chapters the methodology used are described in more details.

2 Aspects of User Satisfaction

Studies have show that satisfaction can be subdivided into five aspects [1]:

- **Efficiency**: this refers to the user feeling that the software is enabling the taks(s) to be performed in a quick, effective and economical manner or, at the opposite extreme, that the software is getting in the way of performance;
- **Affect**: this is a psychological term for emotional feeling. In ths context it referes to the user feeling mentally stimulated and pleasant ot the opposite as a result of interacting with the software;
- **Helpfullness**: this refers to the user's perceptions that the software communicates in a helpful way and assists in the resolution of operational problems;
- **Control**: this sub-scale refers to the feeling that the user has that the software is responding in a normal and consistent way to input and commands;
- **Learnability**: this sub-scale refers to the feeling that the user has that it is relatively straightforward to become familiar with the software and that its tutorial interfaces, handbook etc. are readable and instructive.

3 Software environment

The basic purpose of our usability testing was to evaluate the affects (emotional response), efficiency, learnability, helpfulness, and ease of use of the EducaNext educational portal (Fig. 1). The EducaNext portal addresses the new trends in higher education by providing a web-based tool for the sharing of learning resources. On EducaNext, educators are able to provide learning resources to their peers and specify offer conditions on which interested consumers are required to agree before accessing the learning resources [5, 6].

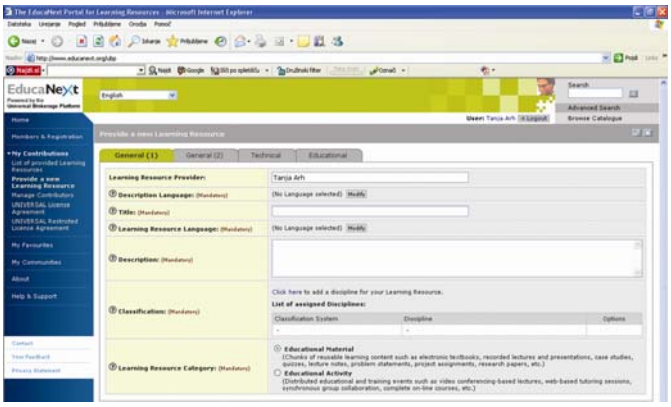

Fig. 1: The educational portal Educanext – v. 1.0

Based on general educational metadata and target-audience specific offer information (e.g. commercial offer, open content-like license agreement, etc.), learning resources are advertised through a catalogue and interest-specific mailing lists. Based on this information, educators can choose and access learning resources from dispersed delivery systems such as video conferencing applications, learning management systems, streaming media servers and standard web servers after agreeing on the terms specified. The process of agreeing on the offer terms is referred to as booking and constitutes an important means for creating awareness about intellectual property rights issues.

For the purpose of SUMI evaluation we used the data of the usability tests on the Slovenian versions of the platform. Standard user test procedures were adopted [2]. Slovenian native speakers were involved.

4 Procedures and observations

Usability testing of EducaNext educational portal was done in computer-equipped rooms with a computer dedicated to each of the participants. The evaluation process was almost identical for every group of participants in Slovenia. The experimenters met with each group for 10 minutes to explain the purpose of the evaluation session and present the methodology of SUMI evaluation. Throughout the detailed explanation about evaluation session, the participants received verbal instructions from the experimenters. The experimenters were present to assist with any difficulties with the questionnaire and to answer questions as they possibly arose. In the second phase, the users were asked to fill the SUMI questionnaire for user-interaction satisfaction. The evaluation sessions lasted about 20 minutes each. During the sessions users were not allowed to ask the evaluator questions.

Three groups of trainees were tested at two different institutions (one academic and the other commercial/governmental) in Slovenia. The first group consisted of 10 trainees, the second of 9, and the third of 12. Age, sex, and employment status were similarly distributed as in the first group that went for general testing. The participants included mostly adults, who were unemployed (60 %), and the rest unemployed. The age range of the participants was 18 to 50. As part of the recruiting process, all participants had some basic computer and web-browser experience. Beyond this basic level, the participants varied in terms of their computer skills as well as in their language skills (mother language and English language).

5 SUMI Questionnaires

The SUMI questionnaire includes 50 questions or statements for which the user selected one of three responses ("agree", "don't know", "disagree"). The statements presented to the participants were addressing their attitudes and their relationship to the system. Once the questionnaires were completed, a dedicated software program called SUMISCO, which comes in the SUMI evaluation package, scored them and compared the results to the standardization database. The mean score of the standardization database is 50, with a standard deviation of 10. Since the standardization database is developed from successful commercial products, a system that achieves a score in the range 40–60 is comparable in terms of usability to most of these products (the standardization database does include score below and above that range).

6 Study Results

The results from the SUMI evaluations are presented in Table 1 in terms of the median, upper, and lower confidence levels. These levels are derived from the global usability scale and each of the five usability sub-scales. The median is the middle score when the scores are arranged in numerical order. It is an indicative sample statistic. The upper and lower confidence limits represent the limits within which the theoretical true score falls 95 % of the time for this sample of users. The standardization database provides, as a general guideline, that a reasonably acceptable commercial software should have a set of sub-scores of 55 or over, and a global score of above 50 with a standard deviation of 10 (with a maximum score of 73). On the global scale, the most reliable of all the SUMI scales indicates that the usability of the evaluated system is comparable to successful commercial systems. In terms of the usability, sub-scales show that the results are consistent and that the scores obtained are above average.

Scale	Lcl	Median	Ucl
GLOBAL n=31	43	56	59
Efficiency	55	58	61
Affect	50	63	66
Helpfulness	57	60	63
Control	33	49	52
Learnability	51	55	59

Table 1: The results of SUMI Questionnaires

The higher values/scores were obtained for Effect, Helpfulness and Efficiency, while the lowest values/score was given to the Control and Learnability.

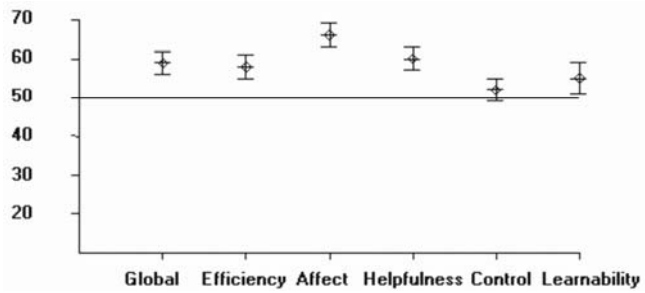

Fig. 2: Comparison of quantitative usability measurements

7 Discussion

Generally speaking, the above findings are consistent with the assumption that users behave rationally when working with an interactive system such as an educational portal EducaNext. The users' behaviors indicate that they were aware of the fact that their own knowledge state slightly deviated from the optimum level required for interacting effectively with the portal. This is an obvious conclusion if the context and the targeted audience, selected according to the set principles, are taken in account. Consequently, the users have shown sufficient engagement regarding exploratory actions. The lower the user's level of domain knowledge required for interacting effectively with the system, the higher the tendency for the user to explore the situation to bridge the knowledge gap. For the provider of the educational service, the most important findings were that the EducaNext portal has shown a high level of learnability, especially in the case of a novice user.

Based on the usability ratings which we gathered for the user interface, as shown in Fig. 2, we received a result that on Global scale the user interface shown better results than average. Also for all other sub-scales was seen that results were better than average and in the desired range of 40 to 60.

In order to improve the score, there was only evidence that the designers need to make modifications in the user interface to improve control and learnability with better navigation and informative functions. Control scores were the most none spreading and this shows that most people did not agree or were undecided about control of the user interface.

There is a need to make faster responds of the software and to make easier path for moving from one task to another task. Also user interface should be more economically in the mean of keystroke using. On the other way it was found that users has to read too much

before starting the portal and do not have a possibility to see all options at a glance.

8 Conclusions

The usability methodology presented in this paper for evaluating the learnability of an educational portal EducaNext is plausible. It provides better understanding of the cognitive mechanism underlying the observed effects and precise information about the tradeoffs in using SUMI methodology.

The results and findings of the study gave important information for the producers and designers of the educational portal EducaNext to know how users learn from their problems in interacting with the system and how effective their workarounds are. This is certainly relevant for the bodies and governmental institutions interested in supporting lifelong learning systems over the Internet and improving the general educational level in the country.

Acknowledgment

This work has been performed in the framework of the European PHARE CBC Slovenia – Austria Grant Scheme S1000-316-01-002, FOCUS-SIAT.

References:

[1] Kirakowski, J., Corbett, M. SUMI: The Software Usability Measurement Inventory. *British Journal of Educational Technology*, 24(3), 1993, pp. 210-212.

[2] Synytsya, K. Learning object metadata: Implementations and open issues - Introduction to the special issue. *Learning Technology Newsletter*, 5(1), 2003.

[3] Dumas, J. & Redish, J. *A Practical Guide to Usability Testing*. Exeter, UK: Intellect, 1993.

[4] ISO. *Guidance on usability specification and measures*. ISO, CD 9241-11, 1992.

[5] Law, L.-C., Maillet, K., Quemada, J., and Simon, B. EducaNext: A Service for Knowledge Sharing, in: *Proceedings of Ariadne Conference*, Leuven, Belgium, November, 2003.

[6] Simon, B., Vrabič G. Resource catalogue design of the UNIVERSAL brokerage platform, *Proceedings of ED-MEDIA,* Tampere, Finland, 2001, pp. 1973–1978.

[7] Nielsen, J. *Usability Engineering*, Morgan Kaufmann, San Francisco, 1993.

[8] Holzinger, A. Usability Engineering for Software Developers. *Communications of the ACM*, 48(1), 2005, pp. 71–74.