A key role of evaluation in human-centered design process: methodologies for authoring shell's usability evaluation

Andrina Granić
Faculty of Natural Sciences, Mathematics and Education
University of Split
Nikole Tesle 12, 21000 Split
Croatia
andrina.granic@pmfst.hr

Vlado Glavinić
Faculty of Electrical Engineering and Computing
University of Zagreb
Unska 3, 10000 Zagreb
Croatia
vlado.glavinic@fer.hr

Abstract – **The discipline of human-computer interaction, HCI, is giving a lot of emphasis on defining methods for ensuring generation of usable interactive computer systems. Redesigning user interface on the basis of iteration of evaluation and design can substantially improve interactive system's usability. Within the research on human-centered interfaces for on-site and for Web-based authoring shells, systems which provide generation of emulators of human teachers in the process of learning and teaching, we consider the development of a suitable usability evaluation methodologies that enable quantification of relevant HCI attributes and successively shells' validation.**

I. INTRODUCTION

In order to design usable interactive systems, usability evaluation plays the fundamental role in human-centered design process. Research in the field of human-computer interaction, HCI, has provided several principles and guidelines that can steer the designers in taking their directions. However, applying good design guidelines by themselves is a good start, but there is no substitute for distinct system evaluation. The most promising approach to the generation of usable systems is the iteration of design and usability evaluation until a satisfactory solution is achieved.

Our research is focused on the creation of usable computerized educational systems like intelligent tutoring systems and their generators, authoring shells. Within the study on usable, human-centered interfaces for on-site along with the one for Web-based educational systems, we consider the development of suitable usability evaluation methodologies that should enable quantification of relevant HCI attributes, and in turn systems' validation and especially identification of inherent weaknesses. Ease of learning, ease of use and general user satisfaction, along with quality and comprehensiveness of content and functional capabilities, will determine the success or failure of the effort. The paper elaborates on relevant issues and methods in addition to usability evaluation methodologies the authors have developed and employed in order to evaluate interface design of an on-site and of a Web-based intelligent tutoring system/authoring shell.

II. HUMAN-COMPUTER INTERACTION AND USER INTERFACES

Methodologies for building usable interactive computer systems have been introduced and refined over the past twenty years under the discipline of *human-computer interaction*, HCI. In order to enhance the quality of the communication between humans and interactive systems, the discipline of human-computer interaction systematically applies knowledge about human purposes along with human capabilities and limitations on the one hand, as well as machine on the other. Human-computer interaction can be defined in many possible ways, e.g. [10], [20], but the crucial point is in considering it as the integrated process of design, implementation and evaluation [9] (see Fig. 1). HCI principles include an early and consistent focus on end users, the tasks they perform and context in which they work, empirical measurements of system usage, as well as iterative development.

Studies show that human-computer interaction/user interfaces is still high on the list of topics with the greatest "knowledge gap", the topic importance mostly exceeding current knowledge [12]. The design and implementation of user interfaces, which can be broadly defined as the two-way communication channel between the human and the functional elements of the machine, is hard and time consuming, because of its complexity and difficulty for implementation, debugging, as well as modification [16]. Consequently, in order to achieve interfaces' almost transparency and enable end users to fully concentrate on the work, the main goal of human-computer interaction as a design discipline is the provision of usable and functional computer systems.

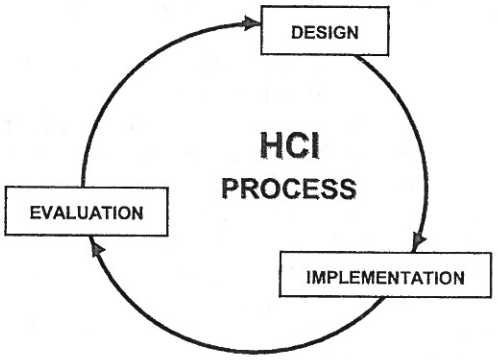

Fig. 1. HCI as the integrated process of design, implementation and evaluation

III. USABILITY AND USABILITY EVALUATION

Usability, as the significant concept in HCI [19], is concerned with making interactive systems easy to learn and easy to use [20], thus providing harmony among four components of any work situation: user, task, environment, as well as the system itself. Studies show that redesigning

a system's interface on the basis of the iteration of design and evaluation can substantially improve usability, e.g. [19], [22], as also seen from Fig. 1. Therefore, in order to create usable interactive systems, it is necessary to augment the standard life cycle to explicitly address usability issues. From the above discussion, it follows that evaluation represents the key phase in human-centered design process for interactive systems. *Usability evaluation* is the process by which the human-computer interaction characteristics of a computer system are measured and weaknesses are identified for correction. Such testing can range from rigorously structured to highly informal, from quite expensive to almost free and from time-consuming to relatively quick. However, while the amount of improvement is related to the effort invested in usability testing, all of these approaches lead to better systems.

Our experience, comprising the design, implementation as well as evaluation of user-centered intelligent computerized educational systems, indicates that extremely useful usability assessment can be performed reasonably easily, quickly and for almost no cost other than employees' time.

IV. METODOLOGIES FOR AUTHORING SHELL'S USABILITY EVALUATION

Educational system, traditionally defined as a community group which includes teachers, as well as students and their joint work in the process of learning and teaching, is nowadays exposed to important changes. Those modifications are also a result of the employment and support of information and communication technology, which enforce the role of computer in education altogether. Our research is focused on user interfaces for computerized educational systems like intelligent tutoring systems and authoring shells.

A. Intelligent Tutoring Systems and Authoring Shells

Intelligent tutoring systems, ITSs, are a generation of computerized educational systems, which attempt to mimic human tutors [5], thus improving the process of learning and teaching. Within this framework special attention is given to accommodate an ITS to the particular domain knowledge, thus leading to the development of *authoring shells*, ASs [1], meant to act as generators of specific ITSs. ASs are intended to adjust to teachers as well as to students within an interactive learning environment by supporting teachers in the development of a series of ITSs for arbitrary domain knowledge and conversely, by enabling students to learn, test themselves and be advised on further work. The literature review admit that communication between users and ITSs/ASs is inherently complex, especially when supporting student interaction because of the student's dealing with concepts (domain knowledge) yet not understood very well [15]. Additionally, poor and unusable interface design can easily prevent students from learning.

Within the field of ITSs and ASs, the latter providing automated ITS generation, we conceptualized, developed and deployed a series of educational systems like authoring shell *Tutor-Expert System*, TEx-Sys [23], Web-based intelligent tutoring system *Distributed Tutor-Expert System*, DTEx-Sys [21] and ITS generator with adaptive interface *Adaptive Knowledge Base Builder*, AKBB [6]. In the following we briefly describe methodologies developed for the usability evaluation of on-site authoring shells - TEx-Sys and AKBB, as well as for the usability assessment of Web-based ITS - DTEx-Sys.

It is a well-known fact that users evaluate the usability of any interactive computer system in terms of quality of relevant user interface. Several different approaches at the operational level to the assessment and measurement of interaction between users and systems are known from the literature, while every one of them considers usability in terms of a number of operational criteria called attributes or parameters. In order to allow usability quantification, those parameters "... *formalize the user behavior to be supported* [and provide] *usability objectives to the level at which it is to be supported*" [4].

B. Methodology for On-site ITS/AS Evaluation

Usability as quality of use in context [3] should be viewed as comprising of two basic aspects: *ease of use* considered primarily as involving subjective judgments and *efficacy in use* considered primarily as involving measures of (human) performance [22]. Consequently, usability evaluation of an on-site ITS/AS interface is based on criteria expressed in terms of:

- objective performance measures of *effectiveness* - the accuracy and completeness with which user achieve specified goals and *efficiency* - the resources expended in relation to the accuracy and completeness of goals achieved - in system's use, as well as
- user's *subjective assessment* of her/his system's usage.

An operational definition of system's usability enables not only setting quantitative goals of execution before the usability evaluation is performed, but also enables the specification of operationally defined criteria for success, because it comprehends following seven measurable attributes: suitability, adaptivity, learnability, memorability, error rate, subjective satisfaction and overall subjective satisfaction. Since user testing, like all empirical studies, requires a theoretical framework for definitions and measures, we have specified usability attributes along with different ways of their measuring according to a formal method for specifying operationally defined criteria for success [25].

Considering different methods of usability evaluation, having in mind that usability can only be meaningfully measured during task performance and that is better to perform any kind of usability assessment than no testing at all, we selected an approach which comprehends formal user testing during users' walkthrough along ITS/AS interface (see Fig. 2. for a snapshoot of an interface), guided with a set of predefined steps. Test users are tested with actual tasks under conditions that are as close to those in the actual usage. *Scenario-based usability evaluation*, as a methodology for an on-site ITS/AS evaluation (for more details see e.g. [8]), is comprehended of walkthrough usability test, memo test and usability questionnaire.

1) Walkthrough usability test: Walkthrough usability test is composed of two parts: (i) scenario-guided task which

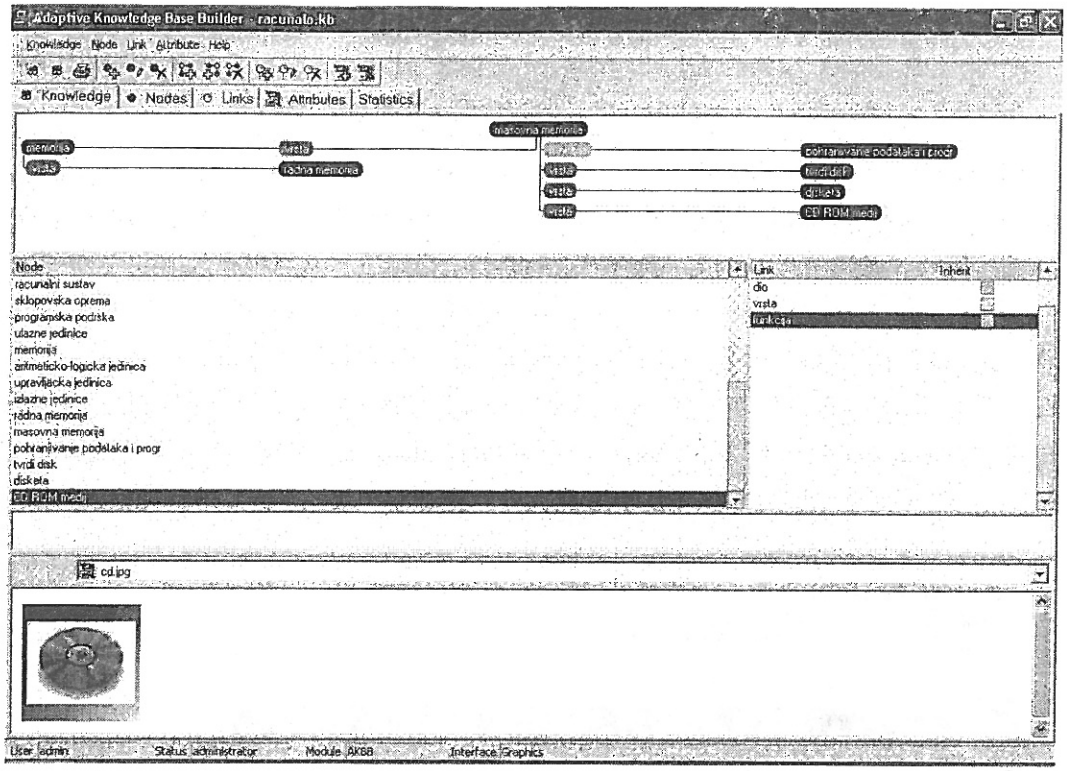

Fig. 2. Snapshoot of *AKBB*, an on-site ITS generator

has to be followed step by step in order to show system's basic functionality as well as main aspects of its interface and (ii) an arbitrary specific task which test users have to perform alone, thus enabling measures for effectiveness and for efficiency of following attributes: suitability, adaptivity, learnability and error rate.

2) Memo test: Memo test is performed after the walkthrough usability test and enables measures for effectiveness and efficiency of the interface memorability attribute by requiring the user to explain the effects of a single command or to write down command name for particular operation.

3) Usability questionnaire: In addition to evaluating "hard" measures like time to complete a task and/or error rates, it is extremely useful to investigate the less observable aspects of interface design that cumulatively contribute to a user's subjective feelings of satisfaction or frustration. Usability questionnaire is due to measurement consistency filled out after the memo test and enables measures for user satisfaction with the system's interface, as well as her/his satisfaction with its suitability, capability of adaptation, ease of learning, memorability and error rate during performance of usability test. Responses are ranked on a 1 to 7 point semantic differential scale. From the standpoint of the single user the responses represent her/his subjective opinion, but as an average value taken from a number of users they indicate an objective value of system pleasantness.

The results obtained through such usability evaluation were subsequently used for determining the interface strength and weakness, hence furnishing a direction in on-site ITS/AS interface design improvement.

C. Methodology for Web-based ITS/AS Evaluation

Although usability engineering has come to play an increasingly important role in conventional interactive system development, it is still rarely part of Web-based system generation despite the fact that there already exist a number of Web style design guidelines, e.g. [24], [13]. On the other hand, employing usability guidelines by themselves does not guarantee the development of usable systems. Usability evaluation, as a distinct validation process, must be performed. In the case of Web-based systems this means taking into consideration similarities and divergences with respect to conventional systems. A Web-based ITS/AS can be considered at the same time a traditional software application and also a Web site, because "*the Web* [and accordingly Web-based ITS/AS] *is delivery medium, content provider and subject matter all in one*" [14]. In addition to, like for any well-designed conventional system or Web site, good usable interface design is crucial. Accordingly, Web-based ITS/AS (see Fig. 3. for a snapshoot of an interface), as a collection of Web pages, must follow an established software development life cycle, beginning with requirements gathering, moving through analysis and design, implementation, testing and finally deployment.

Methodology for Web-based ITS/AS evaluation was derived from the one used to accomplish usability assessment of the on-site version. In order to cope with the above requirements, the methodology is composed of several usability evaluation methods: a scenario-based usability test, a guideline evaluation and a usability questionnaire (for more details see e.g. [7]).

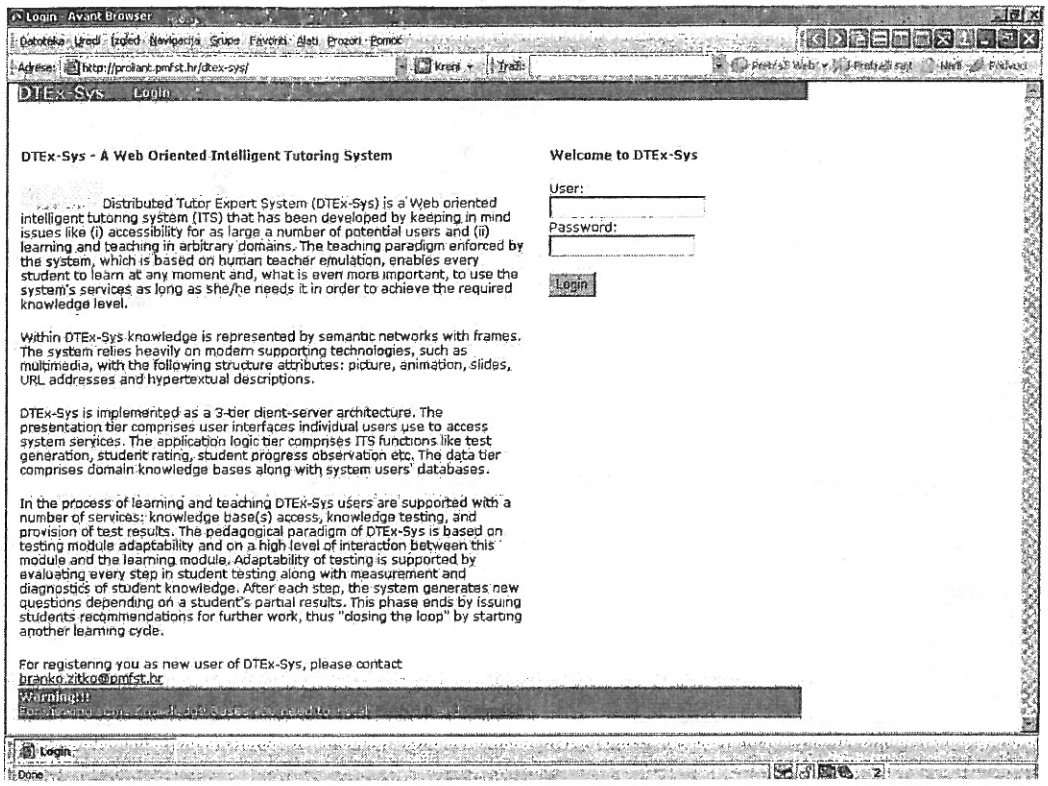

Fig. 3. Snapshoot of *DTEx-Sys*, a Web-based ITS

1) Scenario-based End User Testing: A scenario-based usability test involves representative end-users and scenarios or specific tasks designed to cover the major system's functionality and to simulate expected real life usage patterns. Although measures such as task correctly accomplished or/and task completion time are usually achieved, since we are dealing with educational systems, it is much better to perform some more crucial measurements like how much users actually learn of the information (domain knowledge) contained in the site. Consequently, a scenario-based end user testing enables us to measure: correctness/accuracy of tasks, recognition and recall memory and how much and in what time users actually learn. In order to obtain these measures, a scenario-based usability test comprises following expected usage patterns: (i) three search tasks, (ii) a short examination after some time spent on site and (iii) a quiz enabling an examination of achieved knowledge. The great advantage of such empirical end-user testing is that accomplished results are unquestionable, because user testing will highlight the situations where users actually do have difficulties. When performing scenario-based evaluation we were taking into account the fact that best results come from testing no more than 5 users, because they can find 85% of the usability problems [18].

2) Set of Usability Guidelines: A usage of three to five usability specialists for heuristic evaluation is recommended, because they can usually identify about 75% of the usability problems [17]. However, in order to overcome the problem of not having enough usability specialists that can be involved in evaluation, we decided to perform a "less formal" guideline evaluation, conducted by a group composed of five participants. The Web-based ITS/AS was evaluated using our own set of metrics/guidelines derived through an analysis of the literature on Web-based systems' evaluation, e.g. [2], [13], [11], among which just one set of guidelines (the last one) dealing with interface design evaluation for computer-based educational systems. The set of guidelines was applied as a checklist so the participants had to respond whether the site was to be considered to full-fill the guideline (*'Done'*) or more work was needed (*'Has to be done'*) on a seven-point scale. *'Done'* was scored as 7, so the greater an average on the guideline scale, the better an evaluated aspect of system is.

3) A questionnaire for User-Interaction Satisfaction: Questions in the questionnaire for user satisfaction are formulated according to the ones from the literature. Furthermore, two important facts were also taken into account - it is a design of a Web-based interface and, what is more important, it is an interface of a Web-based educational system that is under evaluation. The most intelligent system in the world does no good if users avoid it because they find it annoying. The usability questionnaire supports valuation and determination of user subjective satisfaction with system's interface, as well as her/his satisfaction with its ease of use, efficiency, likebility, as well as with the attitude the system induces in users during its usage. Participants indicate a level of their agreement with a questionnaire statement on a seven-point Likert scale.

Due to space limitations we are not able to give more details concerning achieved measures and accomplished results when applying both, a methodology for an on-site, as well as for Web-based ITS/AS usability assessment. In brief, measurement of usability attributes through methodology for an on-site evaluation showed some minor user interface weaknesses, hence providing a direction in interface improvement. On the contrary, due to Web-based system's development without an employment of HCI principles and with no evaluation at all, it is not surprising that attributes' assessment with an appropriate developed methodology identified specific problems in several aspects. Consequently, according to achieved results current Web-based version must be and will be redesigned, due to the fact that iteration of design and usability evaluation can substantially improve system's usability. Nevertheless, the fundamental challenge always remains - how to identify usability shortcomings before releasing a new interactive computer system in general (or in the early stages of a redesign), when changes can still be made relatively cheaply and easily.

V. CONCLUSION

In order to enhance the quality of the communication between humans and interactive systems, the discipline of human-computer interaction systematically applies knowledge about human purposes along with human capabilities and limitations on the one hand, as well as machine on the other. Research in the field of HCI has provided numerous principles and guidelines that can steer the designers in taking their directions. However, good design guidelines by themselves do not guarantee generation of usable interactive system. Usability evaluation, as process by which the HCI characteristics of a system are measured and weaknesses are identified for correction, is required. Furthermore, although usability engineering has come to play an increasingly important role in conventional interactive system development, it is still rarely part of Web-based system generation.

Our experience, concerned with the design, implementation as well as evaluation of intelligent computerized educational systems like intelligent tutoring systems and authoring shells, the latter providing automated ITS generation, indicates that extremely useful usability assessment can be performed reasonably easily, quickly and for almost no cost other than employees' time. Within the study on user interface design for on-site, along with the one for Web-based ITS/AS, we consider the development of suitable usability evaluation methodologies that enable systems' validation, quantifying significant HCI attributes. Relevant issues and methods in addition to evaluation methodologies the authors have developed and employed to assess interface design of an on-site, as well as of a Web-based ITS/AS are presented in the paper. Ease of learning, ease of use, as well as general user satisfaction, along with quality and comprehensiveness of content and functional capabilities, will determine the success or failure of such approaches. Obviously usability evaluation, like most methodological process improvements, will gain attention as its benefits emerge through use.

VI. ACKNOWLEDGMENT

This work has been carried out within projects 0036033 *Semantic Web as Information Infrastructure Enabler* and TP-02/0177-01 *Web-based Intelligent Authoring Shell*, both funded by the Ministry of Science and Technology of the Republic of Croatia.

VII. REFERENCES

[1] M. Barton, "Authoring shells for intelligent tutoring system", *7th World Conference on Artificial Intelligence in Education*, Washington, USA, August 16-19, 1995, http://www.pitt.edu/~al/aied/barton.html

[2] N. Bevan, "Usability Issues in Web Site Design", *Proceedings of the HCI International '97*, San Francisco, USA, 24-30 August 1997, Elsevier Press, 1997, also available as Version 3, April 1998, http://www.usability.serco.com/papers/usweb98.pdf

[3] N. Bevan and M. Macleod, "Usability measurement in context", *Behavior & Information Technology*, Vol. 13, No. 1 and 2, 1994, pp. 132-145.

[4] J. Carroll and M. Rosson, "Usability Specifications as a Tool in Iterative Development", in R. Harston (Ed.): *Advances in Human-Computer Interaction 1*, Ablex, Northwood, NJ, 1985, pp. 1-28.

[5] A. Fleischmann, "The electronic teacher: the social impact of Intelligent Tutoring Systems in education", 2000, http://www.student.informatik.tu-darmstadt.de/~andreasf/inhalte/its.html

[6] A. Granić, *Foundation of Adaptive Interfaces for Computerized Educational Systems*. Ph.D. Diss., Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia, 2002, (in Croatian)

[7] A. Granić, V. Glavinić and S. Stankov, "Usability Evaluation Methodology for Web-based Educational Systems", 8th ERCIM Workshop "User Interfaces For All", 28-29 June 2004, Vienna, Austria, accepted for presentation, http://www.ui4all.gr/workshop2004/publications/adjunct-proceedings.html

[8] A. Granić and V. Glavinić, "An approach to usability evaluation of an intelligent tutoring system", N. Mastorakis and V. Kluev (Eds.): *Advances in Multimedia, Video and Signal Processing Systems*, Athens, Greece: WSEAS Press, 2002, pp. 77-82.

[9] S. Greenberg, "Teaching Human-Computer Interaction", ACM Interactions, Vol.3, No. 4, July - August, 1996, pp. 62-76.

[10] T. Hewett, R. Baecker, S. Card, T. Carey, J. Gasen, M. Mantei, G. Perlman, G. Strong and W. Verplank, *ACM SIGCHI Curricula for Human-Computer Interaction*, Report of the ACM SIGCHI Curriculum Development Group, 1992.

[11] M. Jones and J. Okey, "Interface Design for Computer-based Learning Environments", February 21, 1995,
http://ddi.cs.uni-potsdam.de/HyFISCH/Multimedia/Learning/InterfaceDesignJones.htm

[12] T. Lethbridge, "What Knowledge Is Important to a Software Professional?", *IEEE Computer*, May, 2000, pp. 44-50.

[13] M. Levi and F. Conrad, "Usability testing of World Wide Web sites", Research Papers of Bureau of Labor Statistics, 1998,
http://stats.bls.gov/ore/htm_papers/st960150.htm

[14] Th. McManus, "Delivering instruction on the World Wide Web", 1996,
http://www.svsu.edu/~mcmanus/papers/wbi.html

[15] J. Miller, "The Role of Human-Computer Interaction in Intelligent Tutoring Systems", in M. Polson and J. Richardson (Eds.): *Foundations of Intelligent Tutoring Systems*, Lawrence Erlbaum Associates Publishers, Hillsdale, NJ, 1988, pp. 143-189.

[16] B. Myers and M. Rosson, "Survey on User Interface Programming", *CHI'92 Conference Proceedings*, May 3-7, 1992, pp. 195- 202.

[17] J. Nielsen, "How to conduct a Heuristic Evaluation", 2003,
http://www.useit.com/papers/heuristic/heuristic_evaluation.html

[18] J. Nielsen, "Why You Only Need to Test With 5 Users", Alertbox March, 2000,
http://www.useit.com/alertbox/20000319.html

[19] J. Nielsen, *Usability Engineering*, Academic Press, Boston, MA, 1993.

[20] J. Preece, Y. Rogers, H. Sharp, D. Benyon, S. Holland and T. Carey, *Human-Computer Interaction*, Addison-Wesley, Wokingham, England, 1994.

[21] M. Rosić, *Establishing of Distance Education Systems within the Information Infrastructure*, M.Sc. Thesis, Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia, 2000 (in Croatian)

[22] B. Shackel, "Usability – context, framework, design and evaluation" in B. Shackel and S. Richardson (Eds.): *Human Factors for Informatics Usability*, Cambridge University Press, Cambridge, 1991, pp. 21-38.

[23] S. Stankov, *Isomorphic Model of the System as the Basis of Teaching Control Principles in the Intelligent Tutoring System*, Ph.D. Diss., Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, University of Split, Split, 1997 (in Croatian)

[24] B. Tognazzini, "First principles", Nielsen Norman Group, 2003,
http://www.asktog.com/basics/firstPrinciples.html

[25] J. Whiteside, J. Bennett and K. Holtzblatt, "Usability Engineering: Our Experience and Evolution" in M. Helander (Ed.): *Handbook of Human-Computer Interaction*, Elsevier North-Holland, Amsterdam, 1988, pp. 791-817.