Univ Access Inf Soc (2008) 7:209–221
DOI 10.1007/s10209-008-0118-z

LONG PAPER

Experience with usability evaluation of e-learning systems

Andrina Granić

Published online: 13 March 2008

Abstract Progress in the field of e-learning has been slow, with related problems mainly associated with the poor design of e-learning systems. Moreover, because of a depreciated importance of usability, usability studies are not very frequent. This paper reports the experience with the usability assessment of intelligent learning and teaching systems which are based on TEx-Sys model and are intended to enhance the process of knowledge acquisition in daily classroom settings. The applied scenario-based usability evaluation, as a combination of behaviour and opinion based measurements, enabled to quantify usability in terms of users' (teachers' and students') performance and satisfaction. According to the achieved results, the main directions for interface redesign are offered. The acquired experience indicates that useful usability assessments with a significant identification of interface limitations can be performed quite easily and quickly. On the other hand, it raised a series of questions which, in order to be clarified, require further comprehensive research, the more so if the employment of universal design within e-learning context is considered.

Keywords e-Learning system · Intelligent tutoring system · Usability · Scenario-based usability evaluation · Guideline-based evaluation

A. Granić (✉)
Faculty of Science, University of Split,
Nikole Tesle 12, 21000 Split, Croatia
e-mail: andrina.granic@pmfst.hr
URL: www.pmfst.hr/~granic

1 Introduction

In order to achieve maximum interactive system transparency and enable users to fully concentrate on their work, the discipline of human–computer interaction (HCI) systematically applies knowledge about human purposes, human capabilities and limitations on one side, and those about the machine on the other. HCI research has provided numerous principles and guidelines that can steer designers in making their decisions. However, while applying good design guidelines alone is a good start, it is no substitute for a distinct interactive system evaluation. To enable and facilitate the design of usable interactive systems according to usability engineering principles, usability evaluation plays a fundamental role in a human-centred design process.

In today's emerging knowledge-for-all society, knowledge is considered to be a foundation of all aspects of society and economy in general, and the need for its rapid acquisition is more important than ever. E-learning, as an instructional content or learning experience delivered or enabled by electronic technology [33], is placed at the crossroad between information and communication technology (ICT) and education. However, progress in the field of e-learning has been very slow, with related problems mostly associated with the poor interface design of e-learning systems [50]. In order to improve the learning experience and increase the system intelligent behaviour, the findings of relevant studies emphasise the central role of user interface design [7]. Furthermore, despite the important role of usability evaluation, the usability studies in the area of e-learning are not very frequent [44], and a consolidated evaluation methodology for e-learning applications is not available [1, 8].

On the other hand, although computers are being used at different levels of the teaching process (as the subject of

teaching and a tool for supporting the teaching process) and despite decades of research, their use for tutoring (as the teacher itself) in everyday teaching environment has been quite limited [23]. With the intention of accomplishing the latter role, some systems, labelled intelligent tutoring systems (ITSs) [4], heavily rely on artificial intelligence techniques supporting student intelligent guidance. Despite the fact that ITS aim at imitating a human tutor, and their usage for tutoring is still not a part of daily classroom settings, it has been argued that such technology-enhanced learning can improve the overall process of learning and teaching [14]. This viewpoint is supported by research concerning ITSs and their generators—authoring shells (ASs) [27], as exemplified in a series of developed systems based on TEx-Sys model [15, 40, 48]. These systems are partially employed in the process of learning and teaching a number of real subject matters. However, they are still facing problems which need to be considered. Although these systems have enriched tutoring and enhanced the process of knowledge acquisition, there are indications of difficulties in understanding some of their aspects, for both teachers and their students. Most of these difficulties are definitely the result of insufficiently elaborated user interface design and the lack of simplicity and ease in using interaction mechanisms.

This paper reports experience regarding the design and employment of methodologies for the usability evaluation of a series of developed systems for intelligent learning and teaching. A strong motivation to this initiative also came from the reports stating that hardly any usability tests have been going into the design of e-learning technologies, mostly due to failing to perceive the importance of usability, cf. [44]. Apparently, the same problems occur in the research work of the authors. Performed usability studies enabled to become aware of the necessity for ITSs/ASs user interface redesign in order to actually design and deploy useful and usable tools and learning environments. The systems developed so far confirmed that the development was still focusing more on technology than on user-centred aspects. In order to cope with these issues, more or less standard usability assessments and studies have been carried out. Even though the performed studies and achieved evaluation results are encouraging, further research is needed. Significant efforts accomplished in Croatia in both the HCI field and the field of intelligent tutoring just scratched the surface of e-learning "interface issues", and opened a series of questions. In order answer such questions, further extensive research is needed, especially when the application of design-for-all in the e-learning field is considered.

The rest of this paper is structured as follows. Section 2 introduces intelligent tutoring systems, i.e. e-learning systems which emulate a human teacher in an interactive learning environment. The concept of authoring shell, i.e. an ITS generator, is also introduced. Usability issues concerning the e-learning field are presented in Sect. 3. This section also offers a brief introduction to the main issues regarding universal design related to e-learning systems. Methodologies for usability assessment of the systems based on the TEx-Sys model, on-site authoring shells and Web-based intelligent tutoring systems, along with the results obtained from experiments, are discussed in Sect. 4. The same section brings the discussion and interpretation of findings and offers some directions for future research. Finally, Sect. 5 concludes the paper.

2 Intelligent learning and teaching systems

The educational system, traditionally defined as a community group, which includes teachers and students, as well as their cooperative work in the process of learning and teaching, is nowadays exposed to important changes. Such changes are also the result of ICT application and support which enforce the role of computers in education altogether. In such a context, three main roles of computers can be identified: computer as the subject of teaching, computer as a tool in support of teaching process, and computer as the teacher itself [42].

In order to support the latter role in the learning and teaching process, computer-based systems heavily rely on artificial intelligence techniques. This led to the development of intelligent tutoring systems (ITSs), which are e-learning systems intended to support the process of learning and teaching by imitating a human teacher. The emulation process attempts to mimic the behaviour of a human tutor [13], taking into consideration the knowledge about what to teach (domain knowledge), how to teach (teacher knowledge) and whom to teach (student knowledge) [4, 54].

As a need to cover a variety of different knowledge domains has arisen since, instead of having a number of particular ITSs for domains of interest, authoring shells (ASs) [27] have been developed. They are intended to act as generators of specific intelligent tutoring systems in various knowledge domains, as shown in Fig. 1. Authoring shells are meant to accommodate to teachers as well as to students within an interactive learning environment, by supporting teachers in the development of a series of ITSs, and, conversely, by enabling students to learn, test themselves and be advised on further work. It is claimed that, starting from the 1970's when the Scholar system was developed [5], ITSs undoubtedly have improved the process of learning and teaching [14], at the same time taking into account the individuality of a person being taught. In fact, ITSs still represent the best way to enable one-to-one instruction, since they provide the student with a personalised "computer teacher" [39].

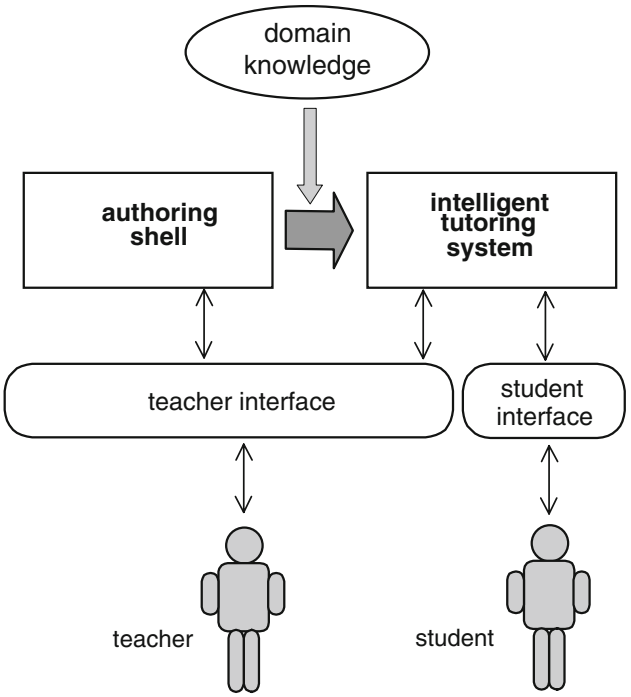

Fig. 1 Authoring shell and intelligent tutoring system

Yet, the progress in the e-learning field has been very slow and ITSs usage for tutoring in the real classroom has been quite limited. It seems that too much of this research is being driven by technical possibilities, while paying inadequate attention to the area of application. The result is an over-ambitious and pre-mature attempt to eliminate the teacher's role in the educational environment [23]. Moreover, huge resources were spent for, e.g. courseware development, and not enough was left to improve the actual quality of acquiring knowledge [28]. These issues have been ignored for quite a long time, hoping that new technologies will somehow resolve the lack of real progress. However, experience has proved so far that these issues cannot be avoided, as they determine the type and scope of tutoring systems that are likely to succeed [35]. Concerning particularly ITSs interaction mechanisms, there persists a general feeling that the development of application technology still has not been accompanied by adequate user interface design. In order to enhance the system ability to improve the learning experience and increase the system intelligent behaviour, recent accounts of ITS architectures attribute a key role to interface design [7, 23], hence avoiding problems related to inadequate or unusable interaction support.

3 Importance of usability evaluation

Research in the HCI field has provided numerous principles and guidelines that can steer designers in taking their decisions. Nevertheless, although applying good design guidelines alone is a good start, it is no substitute for system assessment. To enable and facilitate design according to usability engineering principles, usability evaluation plays a fundamental role in a human-centred design process [18, 30]. Numerous different approaches to the assessment and measurement of interaction between users and systems are known in the literature. Every one of them considers usability in terms of a number of criteria which formalize the user behaviour to be supported, and provides usability objectives at an appropriate level [6]. Moreover, usability as a quality of use in a context [3] should be viewed as comprising two essential aspects: (1) efficacy in use, considered primarily as involving measures of user performance, and (2) ease of use, considered primarily as involving subjective judgements [41].

Although usability is the basic parameter for the evaluation of e-learning technologies and systems [51], the idea of e-learning usability is still quite new [37]. Concerning usability evaluation methods, it has been claimed that usability assessment needs further consideration of the learning perspective [31]. There are some approaches adapted to e-learning [46], although some authors propose applying heuristics without further adjustment to the e-learning context [10, 19, 34]. Obviously, there is a need for further research and empirical evaluation [47], since an established set of heuristics [52] and a joint evaluation methodology for e-learning applications are not yet available [1, 8].

Furthermore, the employment of design-for-all in e-learning environments promotes individualization and end-user acceptability, ensuring that usability and accessibility should be design concerns. For this reason, the design of accessible and easy to use e-learning system able to address the needs of all potential users requires additional considerations. The main issues regarding universal design related to e-learning systems include:

- *Learner-centred design paradigm:* the same practices followed by the HCI community must be used in order to ensure learnability, a major issue for e-learning, as rephrased by Don Norman [12].
- *Context of use approach:* in order to match users' needs in the natural working environment, e-learning system should be seen in terms of a four-component model of HCI [36]: whether the user for whom it is designed can use it with acceptable levels of usability and accessibility, for the tasks that s/he needs to do, in the local environment in which these tasks take place, using the available technologies.
- *Individualized approach:* the consideration of users' different individual characteristics relevant to learning

styles and preferences fosters individualization and end-user acceptability.
- *Pedagogical framework:* the support of (new) pedagogical approaches that blend new and old ways of learning in order to maximise the learning potential of technology.
- *Guideline framework:* the employment of usability and accessibility guidelines for e-learning quality assessment.

Comprehensive research concerning an evaluation methodology addressing such a broad spectrum of issues is needed. The assessment presented in this paper concentrates on the system usability aspects. The usability evaluation techniques employed in the assessment of on-site and web-based systems for intelligent learning and teaching, as well as the evaluation procedure and the results achieved, are described hereinafter.

4 Evaluation of interaction design in intelligent learning and teaching systems

As mentioned earlier, usability studies in the e-learning field are not very frequent, despite the important role that usability plays in the success of every e-learning system. If the system's interface is not transparent and easy to use, the students are concentrated on interaction aspects and not on acquiring specific domain knowledge. This applies in particular to systems, which emulate the human teacher in the process of teaching, whose usage for tutoring in the actual learning environment in general has been quite restricted and inadequate. In order to enhance learning and teaching process, a series of systems based on the Tutor-Expert System (TEx-Sys) model [48] has been designed, implemented and deployed at the University of Split in Croatia. However, due to development process, which did not embrace usability design principles and guidelines, nor did apply any usability assessment at all, the resulting interaction design turned out to be an obstacle to comfortable and efficient usage. Consequently, the necessity emerged of a comprehensive usability study and the employment of an adequate evaluation methodology in order to redesign the user interfaces of the developed systems.

While considering different methods of usability evaluation, taking into account that usability can be significantly quantified during task performance, an approach was selected which consists of the user walk-throughs within system interface, guided by predefined steps. Due to the fact that usability, as a quality of use in context, is related to the process of use, the steps the user must take and how (s)he achieves the results, the usability evaluation of the systems developed for intelligent learning and teaching is accomplished by testing usability with real users. It is based on criteria expressed in terms of [21, 26]:

- objective performance measurement of *effectiveness* (the level of correctness and completeness with which users achieve specified goals) and *efficiency* (the resources expended in relation to the correctness and completeness of goals achieved) in using the system, and
- users' *subjective assessment* of the system usage.

These objectives stress the need to quantify usability in terms of user performance and satisfaction. They are measured by the extent to which the intended goals of use are achieved, the resources that have to be expended in order to achieve them, and the extent to which the user finds the use of the e-learning system acceptable. Such testing is understood to be a combination of behaviour and opinion based measures with some amount of experimental control, usually chosen by an expert. It affords information about how users (teachers or students) use the system (authoring shell or intelligent tutoring system, respectively) and identifies the exact problems with a particular interface being evaluated.

In the light of these considerations, the evaluation methodology employed in the assessment of operational systems for intelligent learning and teaching based on TEx-Sys model, as well as the related evaluation procedures along with the results achieved, are described in the following chapter.

4.1 Usability assessment of an on-site e-learning system

With the intention of creating a learning and teaching environment suitable for students' individual requirements, two on-site e-learning systems were designed and developed: (1) an intelligent hypermedia authoring shell, the TEx-Sys, which provides the means for developing specialized ITSs for particular domains of education [48], and (2), an arbitrary domain knowledge generator with adaptive interface, the adaptive knowledge base builder (AKBB) [15], built on TEx-Sys basic functionality. As regards design, TEx-Sys has a static user interface, while AKBB enforces a simple adaptive mechanism, which selects the most appropriate interface out of a number of options according to the run-time tracing of user behaviour. Presently, the adaptive nature of the AKBB user interface is related to interaction style adaptivity rather than to the educational aspects of interface. Within an interactive learning and teaching environment, both TEx-Sys and AKBB support teachers in the development of a series of intelligent tutoring systems for an arbitrary subject matter.

However, with regard to their design, it is important to point out that both were developed without considering HCI principles of usable design and without applying any usability evaluation whatsoever.

Before introducing the evaluation methodology applied for the system assessment, a brief insight into their design in general is provided by means of a short description of the AKBB's interface. The interactive window of the AKBB system is divided into three sections (see Fig. 2 for the screenshot of its interface). The uppermost section displays the structure of specific domain knowledge represented by semantic networks. Namely, in the learning and teaching systems knowledge representation is based on semantic networks, because of their ability to express the cognitive model of human memory and reasoning [32]. Basic elements of domain knowledge are nodes and links [38]. Nodes, denoting the domain knowledge objects, are semantically connected with links. Figure 2 depicts a part of selected domain knowledge related to computer hardware. The active (selected) node from domain knowledge, i.e., the node "memory", is displayed in the centre, its parent node "hardware" on its left and its children on the right. A semantically correct interpretation of the connections between the active node "memory" and the rest of the nodes is as follows: random access memory is a kind of memory; read only memory is a kind of memory; memory is a part of hardware; input device is a part of hardware and so on. In order to refine and enrich the knowledge object description, some documents and/or media files can be attached to each node (for example pictures, animations, URL addresses, hypertextual descriptions and alike). They are displayed in the middle section of the interactive window. In the above example, just a picture "memory.jpg" of the selected node "memory" is shown. The central section also contains the list of domain knowledge nodes (left side of the section) and the list of employed links (section right side). The window bottom part includes command line and displays the related interaction history.

4.1.1 *Applying usability evaluation techniques*

Taking into account that the usability of a particular system depends on the characteristics of the users, the tasks and the system purpose, there is no simple definition or meaningful single measure of usability. Moreover, usability indicators (attributes) take on empirical values during performance and can be evaluated only during use of the system. Consequently, in order to understand the effect of TEx-Sys and AKBB design in a sample work situation, a detailed work scenario was elaborated, consisting of a sequence of typical system tasks and user actions. Usability goals will be achieved if the system potential is actually used both effectively (to a specified level of user performance) and happily (to a specified level of subjective assessment). Hence, the scenario-based usability evaluation comprises:

- walkthrough usability test,
- memo test, and
- usability satisfaction questionnaire.

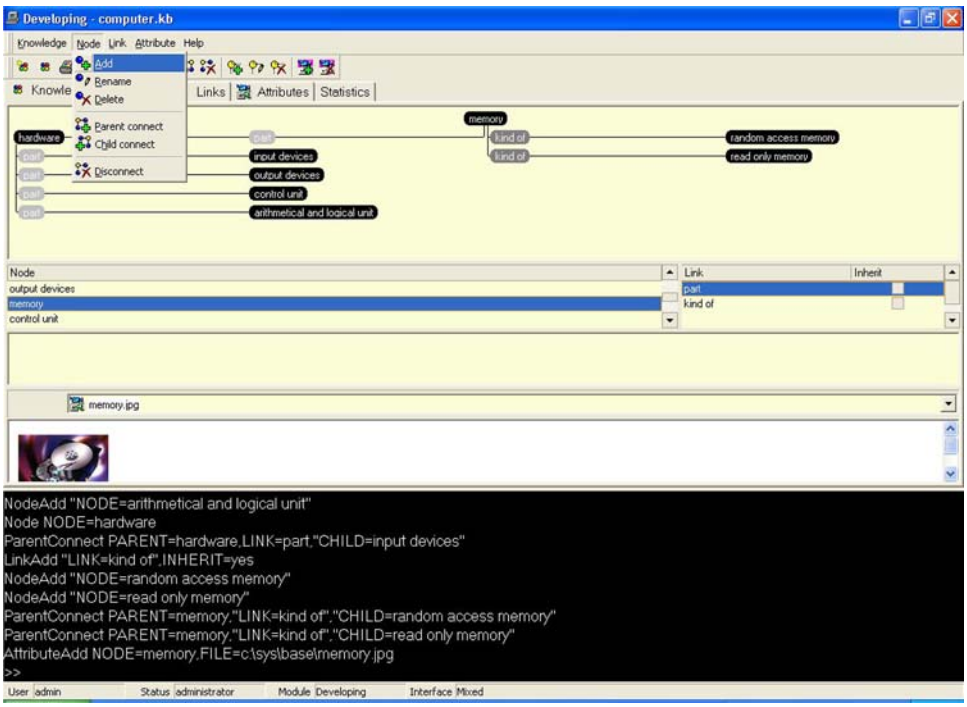

Fig. 2 Screenshot of AKBB

The walkthrough usability test is composed of two parts: (1) a scenario-guided task of a particular domain knowledge generation to show the system basic functionality and main aspects of its interface, along with (2) a specific task related to a particular knowledge generation, which test users (here involving teachers) have to perform by themselves. Accordingly, the walkthrough end-user testing enables to determine the following attribute values:

- *suitability*, expressing the degree of appropriateness of the system to the tasks which have to be accomplished,
- *learnability*, which conveys how easy it is for the user to learn the system and how rapidly (s)he can begin to work with it, and
- *error rate*, which reflects the error ratio while working with the system.

A memo test is performed subsequently to the walkthrough test and enables the measurement of interface memorability by requiring a user to explain the effects of a single command or to write down the command name for a particular operation. In addition to acquiring "hard" measures, like the time needed to complete a task and error rates, it is extremely useful to investigate the less observable aspects of interface design that cumulatively contribute to the user's subjective feelings of satisfaction or frustration. The most intelligent system is not useful if users avoid it because they find it annoying. Consequently, a usability satisfaction questionnaire is filled after the memo test and enables the measurement of users' subjective satisfaction with interface aspects. Questions in the usability questionnaire are formulated according to those found in reference literature, cf. [16, 17, 25, 43]. Responses are ranked on a seven-point semantic differential scale. From the standpoint of a 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 the system pleasantness.

4.1.2 Evaluation session and experimental results

Test users, as representative as possible of the targeted users of TEx-Sys and AKBB system (teachers instructing chosen subject matter) were tested with actual tasks (knowledge base generation concerning selected topic), under conditions as close as possible to those of actual system usage (classroom environment). When performing usability evaluation, the results of a new empirical research were taken into account [11], showing benefits of increased five-user sample size in usability testing [29]. The assessment process was carried out with a group of ten graduate students with a teacher diploma in Computer Science and Mathematics. The experimenter met with the group in order to explain the purpose of the evaluation and present the overview of the usability assessment. At the end of this initial meeting, the evaluation was performed.

This assessment session, comprising a walkthrough usability test, a memo test and a questionnaire for user-interaction satisfaction, enabled the acquisition of main measurements. Since user testing, like all empirical studies, requires a theoretical framework for definitions and measures, quantifiable attributes were defined as shown in Table 1, cf. [21, 26]. In order to assess whether the interface design meets the set specifications and to what extent, all usability attributes were specified by seven items according to a formal method for the definition of success criteria, cf. [41]. It was not easy for the developers to predict what the users (in the specific case, teachers) really wanted and how they would react to the system interface. Thus, the first step of assessment was to work with (potential) users in order to establish system goals. With the intention of specifying values for the worst case and indirectly for the planned level as well as for the best case, pilot testing was performed.

The measurement results obtained through the scenario-based usability evaluation of TEx-Sys authoring shell are

Table 1 Usability attributes specification of an on-site authoring shell

	Effectiveness measurements	Efficiency measurements	Satisfaction measurement
Suitability	Percentage of goals achieved during the walkthrough test	Time to complete the walkthrough test	Rating scale in the questionnaire
Learnability	Percentage of goals achieved when completing a task from the walkthrough test	Time to complete a task from the walkthrough test	Rating scale in the questionnaire
Error rate	Ratio between errors and performed actions when completing a task from the walkthrough test	Number of errors when completing a task from the walkthrough test	Rating scale in the questionnaire
Memorability	Percentage of functions memorized successfully during the memo test	Time to complete the memo test	Rating scale in the questionnaire
Subjective satisfaction			Rating scale in the questionnaire
Overall subjective satisfaction			Rating scale in the questionnaire

reported in Table 2. Although the measures achieved for all attributes were better than the specified worst case, and almost all quite close to the planned value, the debriefing session confirmed the need for interface redesign. Therefore, in order to acquire more complete feedback from the intended users, the quantitative task-based evaluation approach was enhanced with qualitative evaluation methods, observations and interviews.

The teachers found the presentation of domain knowledge slightly confusing and preventing them to quickly grasp the knowledge structure. Such presentation is the "reflection" of TEx-Sys as well as AKBB knowledge representation based on a semantic network. Obviously, it does not convey in a transparent way the semantics of the connected domain knowledge objects, and thus it impedes users in getting clear and unambiguous view of a particular subject matter (see for example the screenshot of the AKBB interface in Fig. 2).

Summarizing, the results of usability measurements along with the teachers' comments collected after testing indicated the need for a number of improvements. These include:

- introducing different ways to create relationships between domain knowledge objects, resulting in more usable interaction and in faster creation of the knowledge base,
- hiding the internal structure of the domain knowledge base, and
- providing a richer and semantically more meaningful set of widgets.

4.2 Usability assessment of a web-based e-learning system

Distributed tutor-expert system (DTEx-Sys) is a web-based intelligent tutoring system which, within an interactive learning environment, enables students to learn diverse domain knowledge, test themselves, be advised on further work, consult a teacher and access help when needed [40] (see Fig. 3). The approach to its usability evaluation is derived from the one used to carry out both TEx-Sys and AKBB assessments, and is based on the combination of behaviour measures and subjective opinions.

4.2.1 *Enhancement of user testing with guideline-based evaluation*

The evaluation was based on criteria expressed in terms of the objective performance measures in using the system, as well as in terms of users' subjective assessment. Additionally, the evaluation methodology employed in both TEx-Sys and AKBB assessment was enhanced by combining user testing with a kind of "less formal" heuristic evaluation, i.e. guideline-based evaluation. A major strength of such an approach is the possibility to supplement results from both the guideline-based evaluation and the empirical user-based one, enhanced by users' feedback on their comfort while working with the system. Consequently, the DTEx-Sys usability assessment includes:

- scenario-based end-user testing,
- usability satisfaction questionnaire, and
- guideline-based evaluation performed by means of a set of metrics/guidelines.

A scenario-based usability test involves representative end-users (in the specific case, students) and scenarios designed to cover the major system functionality and to simulate the expected real life usage patterns when learning and testing the specific domain knowledge. Although measures such as correct task accomplishment and time are usually performed, other types of measurements are also crucial, such as how much students actually learn of the information (in fact domain knowledge) offered by ITS. Consequently, a scenario-based end user testing enables to measure

- task correctness,
- recognition and recall memory, and
- how much and in what time students actually learn, expressed through the *achieved knowledge* and *quiz solving time* attributes, respectively.

Due to the fact that communication between students and ITSs is inherently complex, because of students' dealing

Table 2 TEx-Sys usability measurement results

	Planned level	Current level		
		Effectiveness measurements	Efficiency measurements	Satisfaction measurement
Suitability	90%/45 min/5.5	86%	49 min	4.4
Learnability	85%/20 min/6	78%	27 min	5.5
Error rate	0.05/1/5	0.04	1.6	4.7
Memorability	80%/5 min/5.5	68%	6.6 min	4.7
Subjective satisfaction	−/−/6	–	–	5.0
Overall subjective satisfaction	−/−/6	–	–	4.8

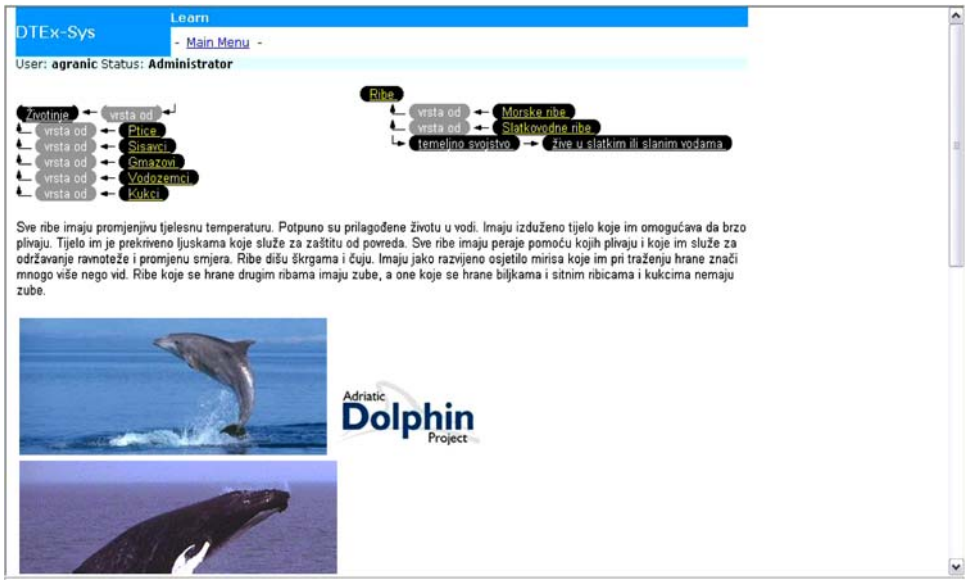

Fig. 3 Screenshot of DTEx-Sys

with the concepts (domain knowledge) they still have not fully grasped, the very last measures are quite significant. In order to obtain these measures, a scenario-based usability test comprises: (1) three search tasks, (2) a short examination after some time spent working on the system, and (3) a quiz to test the acquired knowledge. The great advantage of such empirical end-user testing is that the achieved results are unquestionable. Unlike heuristic evaluation, where HCI experts speculate as to what may cause the users difficulties, this end-user testing highlights the aspects where users actually do have difficulties.

Besides obtaining measures such as the time to complete a task and error rates, it is extremely useful to obtain users' opinions. A *usability questionnaire* supports the assessment of students' *subjective satisfaction* with ITS interface, their satisfaction with its ease of use, efficiency, likeability and with the attitude that the system induces in them during its usage. Users indicate a level of their agreement with a questionnaire statement on a 7-point Likert scale. Questions in the questionnaire for user satisfaction are formulated according to those found in reference literature, *cf.* [16, 43, 53]. Furthermore, two important facts are also taken into account, namely that it is the design of a web-based interface and, what is more important, the interface of a web-based e-learning system that is under evaluation.

In order to overcome the problem of not having enough usability experts who could be involved in the DTEx-Sys evaluation process, *guideline-based evaluation* was performed by "instant experts" [55], i.e. computer scientists who learnt the evaluation methodologies and applied them successfully. A set of guidelines, derived through a literature overview (e.g. [2, 22, 24]) was adjusted to the e-learning context. Experts had to comment on whether the system was considered to have followed the guideline or not. Namely, the majority of authors do not suggest assigning marks to the set of their guidelines; hence the assignment of scores is strictly arbitrary. However, when coupled with the scenario-based end user testing and the questionnaire for user interaction satisfaction, it should nevertheless provide an accurate and objective usability assessment.

4.2.2 Assessment procedure and achieved results

The evaluation was carried out by one experimenter who met with participants to explain the purpose of sessions and to present the assessment overview. In order to get ready for the evaluation session, the experimenter prepared several documents: (1) an evaluation process overview, describing the objectives and target audience, (2) the expected DTEx-Sys usage patterns/tasks along with search tasks and a short exam, (3) a questionnaire for user-interaction satisfaction, and (4) a list of usability principles/guidelines.

The assessment process was carried out in parallel with two separate groups. The first was a group of three instant experts involved in guideline-based evaluation. While inspecting DTEx-Sys interface, experts were identifying potential usability problems, linking each problem found to the specific guideline it violated. A set of guidelines was used as a checklist so that experts had to respond whether the evaluated system was considered to have followed the guideline ("Done") or more work was needed ("Had to be done") on a seven-point scale. "Done" was scored as seven, so the greater the average on the guideline scale, the better the evaluated aspect of the system.

The second evaluation group consisted of ten test users, representing as much as possible the target DTEx-Sys users

(students learning the selected subject matter). The group involved five students of Computer Science and Mathematics and five students of Computer Science and Polytechnics, both attending the second academic year. When performing a scenario-based walkthrough, a participant started at DTEx-Sys home page. The first three tasks concerned searching for the specific information related to the selected domain knowledge, without using a search tool or the "Find" command. Subsequently, participants were instructed to spend half-an-hour learning as much as possible about the specific subject matter. This was a preparation for a short examination and for a final task—answering questions related to the acquired knowledge. For each student, quiz questions were individually generated by DTEx-Sys itself. Finally, participants were asked to fill in the usability questionnaire, indicating the level of their agreement with a statement on a seven-point Likert scale. Main measurements from the scenario-based usability testing and from the questionnaire on user-interaction satisfaction are presented in Table 3. In order to specify values for the planned level, a pilot testing was performed.

Since DTEx-Sys users were students (learners), the designed interfaces do not need support "doing tasks", but support "learning while doing tasks", cf. [20]. As evident from performed measurements, the time to complete a task was not measure, because other measurements were considered as more fundamental, such as:

- how much students actually learn of the provided domain knowledge offered by the intelligent tutoring system itself—*achieved knowledge* (information obtained from a scenario-based end user testing), and
- do students feel efficient while working with the system—*efficiency* aspect from *subjective satisfaction* (information obtained from a questionnaire for user-interaction satisfaction).

Since the system was developed without employing any usability evaluation, it is not surprising that instant experts identified specific problems at almost all levels. While inspecting the DTEx-Sys interface they found various design problems and noticed that more work was needed with respect to the following:

- design of an effective home page that will establish the system identity and give a clear overview of the content,
- provision of links on each page to a list of local content, a site map and the home page,
- the provision of "visual effects" to offer students a visual feedback and information as to where they are and to where they can go, and
- design which will make actions and options visible and which will not rely on the user remembering information.

The violation of the last guideline can also be recognized in the score for one of the five major measures acquired from a scenario-based end user testing—*memory*. It is obvious that the current DTEx-Sys design relies on the user remembering and recalling information and not, as it should, on recognizing it.

4.3 Discussion and future work

Based on performed usability studies and related measurements, the same conclusion arises when evaluating TEx-Sys and AKBB, which are on-site e-learning systems, as well as DTEx-Sys, which is a web-based system, namely that their user interface needs redesign. The summarised results of usability assessment, along with the users' (teachers and students) comments and observations collected upon performing evaluation sessions, indicate the necessity of a number of improvements in each and every design.

Yet, it is strongly believed that the presentation of domain knowledge itself is to be accounted for most of the required improvements. Both teachers and students found the arrangement of domain knowledge somehow confusing, thus preventing them to quickly understand the domain

Table 3 DTEx-Sys measurement results obtained from user testing

	Measuring method	Planned level	Current level
Task correctness	Percentage score based on the number of correct answers in three search tasks	90%	79%
Memory	Percentage score based on no. of correct answers to multiple-choice questions and no. of correctly recalled answers	80%	55%
Achieved knowledge	Score obtained after performing the quiz on specific subject matter generated for each student by DTEx-Sys itself	58/58	46/58
Quiz solving time	Time spent in quiz solving	5 min	6.4 min
Subjective satisfaction	Rating scale in the questionnaire	6	4.2

knowledge structure. Such arrangement is a "reflection" of the knowledge presentation based on semantic networks. Specifically, in evaluated e-learning systems, as well as in the entire series of the developed systems founded on the TEx-Sys model, knowledge presentation is based on semantic networks, with nodes and links as main components. Nodes, denoting domain knowledge objects, are semantically connected through links. A semantically correct interpretation of connections between the active/selected node of domain knowledge and its closest neighbour nodes (parent node and children nodes) is already explained in previous sections (see also the system screenshots in Figs. 2 and 3).

According to the obtained usability evaluation results and the users' comments, the presentation of domain knowledge itself was the main obstacle for acquiring the planned values of measured attributes in the performed assessments. Obviously, this presentation fails to convey in a transparent way the semantics of the connected domain knowledge objects, thus impeding users in obtaining a clear and unambiguous view of a particular subject matter.

In order to hide as much as possible the internal structure of the domain knowledge base, the knowledge presentation was redesigned in both the on-site and the web-based e-learning system by introducing a new way of creating relationships between knowledge components (see Fig. 4 as an illustration of better presentation of nodes and links from domain knowledge). This will hopefully result in:

- more transparent and intuitive, i.e. usable interaction for teachers and students while interacting with authoring shells and intelligent tutoring systems (respectively),
- faster creation of a specific knowledge base resulting in specialized ITS for a particular domain of education (thus targeting the teacher population), and
- faster learning of a specific subject matter using specialized ITS (thus targeting the student population).

Furthermore, according to the usability study results, the new "look and feel" of the redesigned interface of the web-based e-learning system will encompass:

- a well-organized home page that will (1) establish the e-learning system identity, and (2) structure the provided information hierarchically in order to become meaningful to the user (a new way of presenting relationships between domain knowledge components will further enhance this issue),
- information which will let users know where they are and where they can go; this is very important for students while learning new concepts and subject matters,
- a consistent page layout of the e-learning system which will indicate similar concepts through identical terminology and use of graphical expressions; this is also vital while learning new domain knowledge, and
- visible options in order to determine which actions are possible at any moment; this implies design for recognition rather than recall.

The studies and assessment results presented in this paper report experience regarding the usability aspects of the systems developed for intelligent learning and teaching. On the other hand, the objective of universal design is to provide accessible and usable IST products and services, raising the so-called "good user-centred design" paradigm to a more encompassing concept of addressing the needs of all potential users, cf. [49]. Therefore, the application of design-for-all in e-learning environment encourages end-users' acceptability and personalization, hence avoiding the need for later adaptations. Unfortunately, studies have regularly shown that the accessibility of web-based systems in general falls short of an acceptable level [45] and most of existing efforts related to accessible e-learning propose guidelines that primarily address technical accessibility issues [9]. When considering e-learning systems, it has been argued that usability evaluation needs further consideration of the learning perspective, e.g., [31, 46], although some authors propose applying heuristics without further adjustment to the e-learning context, e.g. [10, 34]. Nevertheless, since in e-learning interfaces are needed that support "learning while doing tasks", there should be a synergy between the process of learning and a user's (student's) interaction with the system, hence taking into account the different ways users learn and ensuring their natural and flexible interactions as well.

Consequently, when designing an accessible and easy to use e-learning system, system able to address the needs of all potential users, it is important to consider the key issues that include learner-centred design paradigm, context of use approach, individualized approach, pedagogical framework as well as guideline framework. The issues concerning the universal design within e-learning context and the consequential enhanced evaluation methodology require methodical research that will surely result in a broad spectrum of issues and questions.

5 Conclusion

Current efforts in designing the IST products and services that satisfy the needs of all potential users of today's emerging knowledge-for-all society address user interfaces as one of key issues. They place the individual at the very heart of development, emphasizing the necessity

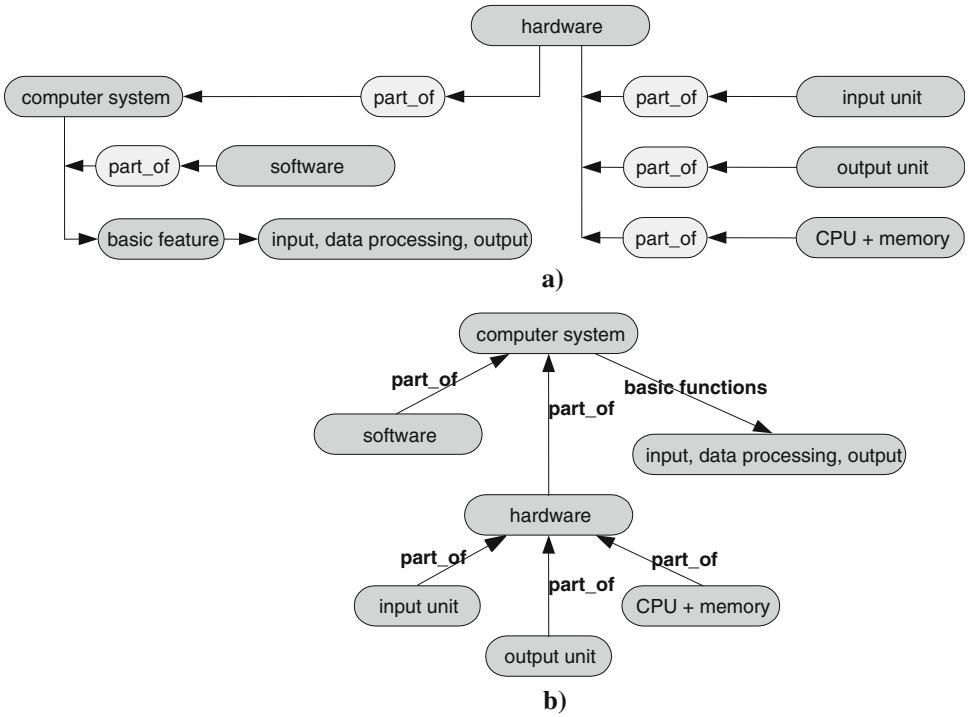

Fig. 4 Domain knowledge presentation with selected node *hardware*: **a** current interface design, **b** look and feel of the future redesigned interface

to design technologies for users and to make interaction accessible and usable. Knowledge is the most important resource in the context of the knowledge-for-all society, and the need for its rapid acquisition is more important then ever. Within this framework, e-learning systems have a particularly important role, because of the increasing need to support educational flexibility as well as self-education and life-long learning. Of the entire area of computer-based educational systems, increased is put on systems that allow for student intelligent guidance and that are as labelled intelligent tutoring systems (ITSs). In order to deal with a variety of different subject matters, authoring shells (ASs) have been conceptualised for the development of specialized ITSs for a particular domain of education.

Despite the fact that ITS's imitation of a human teacher and its usage for tutoring is still not a part of everyday teaching environment, it is claimed that such technology enhanced learning has advanced the process of learning and teaching. This viewpoint is supported by research concerning ITSs and ASs, as exemplified in a series of developed systems based on TEx-Sys model. This paper reports the experience regarding the design and employment of methodologies for usability evaluation of a number of developed systems. Their moderate usage in daily classroom settings has confirmed that so far the development focus was more on technology aspects rather than on user-centred design issues. In order to deal with these issues, a relatively simple usability assessments and studies have been performed. Usability goals will be achieved if the system capability is actually used to a specified level of user performance, as well as to a specified level of subjective assessment. These objectives enabled to highlight the need of quantifying usability in terms of user (teachers and students) performance and satisfaction.

The scenario-based usability evaluation, used for the assessment of the TEx-Sys and AKBB operational on-site authoring shells, engaging teachers as test users, consists of: (1) a walkthrough usability test, (2) a memo test, and (3) a usability satisfaction questionnaire. Furthermore, in order to assess the DTEx-Sys operational Web-based intelligent tutoring system, the above methodology was enhanced with a kind of "less formal" heuristic (i.e. guideline-based) evaluation. This enabled to supplement the results from empirical user-based assessment and the users' (students) feedback on their comfort while working with the system with the guideline-based evaluation performed by "instant experts". Consequently, the obtained measurement results and debriefing sessions indicated the necessity for a number design improvements.

It is argued that the presentation of domain knowledge, which both teachers and students find confusing, accounts for most of the required improvements. The arrangement does not convey in a transparent way the semantics of the connected domain knowledge objects, thus impeding users in obtaining a clear and unambiguous view of a particular subject matter. The new "look and feel" of the redesigned interface of the systems based on TEx-Sys model aims at hiding as much as possible the internal structure of the knowledge base, thus targeting the improvement of

learning experience and also the enhancement of the system intelligent behaviour.

Overall, all the research work initiated and undertaken by the authors of this paper in the HCI field and the field of intelligent tutoring prompted the e-learning "interface issue" and opened a series of questions. In order to provide answers, further thorough and systematic investigation is needed, especially when the application of design-for-all in the field of e-learning is considered.

Acknowledgments Author would like to acknowledge support for the work presented from the project 177-0361994-1998 *Usability and Adaptivity of Interfaces for Intelligent Authoring Shells* funded by the Ministry of Science, Education and Sports of the Republic of Croatia.

References

1. Ardito, C., Costabile M.F., De Marsico, M., Lanzilotti, R., Levialdi, S., Roselli, T., Rossano, V.: An approach to usability evaluation of e-learning applications. Univ. Access Inf. Soc. **4**, 270–283 (2006)
2. Bevan, N.: Usability issues in Web Site Design. In: Proceedings of the HCI international '97. San Francisco, USA, 24–30 August 1997. Elsevier Press, Amsterdam (1997)
3. Bevan, N., Macleod M.: Usability measurement in context. Behav. Inf. Technol. **13**, 132–145 (1994)
4. Burns, H., Capps, C.: Foundations of intelligent tutoring systems: an introduction. In: Polson M., Richardson J. (eds.) Foundations of Intelligent Tutoring Systems, pp. 1–18. Lawrence Erlbaum Associates Publishers, Mahwah (1988)
5. Carbonel, R.: AI in CAI: an artificial-intelligence approach to computer-assisted instruction. IEEE Trans. Man Mach. Syst. MMS 11(4), 190–202 (1970)
6. Carroll, J., Rosson, M.: Usability specifications as a tool in iterative development. In: Harston R. (eds.) Advances in Human–computer Interaction 1. Ablex, Northwood, pp 1–28 (1985)
7. Collins, A., Neville, P., Bielaczyc, K.: The role of different media in designing learning environments. Int. J. Artif. Intell. Educ. **11**, 144–162 (2000)
8. Costabile, M., Marisco, M., Lanzilotti, R., Plantamura, V., Roselli, T.: On the usability evaluation of E-learning applications. In: Proceedings of the 38th Annual Hawaii International Conference on System Sciences (HICSS'05)—Track 1, vol. 01. 6.2 (2005)
9. De Marisco, M., Kimani, S., Mirabella, V., Norman, K.L., Catarci, T.: A proposal toward the development of accessible e-learning content by human involvement. Univ. Access. Inf. Soc. **5**, 150–169 (2006)
10. Dringus, L.: An iterative usability evaluation procedure for interactive online courses. J. Int. Instr. Dev. **7**, 10–14 (1995)
11. Faulkner, L.: Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav. Res. Methods Instrum. Comput. **35**, 379–383 (2003)
12. Feldstein, M.: What is ''Usable'' e-Learning? eLearn Magazine. ACM Press, New York. http://www.elearnmag.org (2002)
13. Fleischmann, A.: The electronic teacher: The social impact of intelligent tutoring systems in education. http://www4.in.tum.de/~fleischa/papers/its.html (2000)
14. Fletcher J.: Evidence for learning from technology—assisted instruction. In: O'Neil, H., Perez, R. (eds.) Technology Applications in Education: A Learning View, pp. 79–99. Lawrence Erlbaum Associates, New Jersey (2003)
15. Granić, A.: Foundation of adaptive interfaces for computerized educational systems. Ph.D. Dissertation. Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia (in Croatian) (2002)
16. Harper, B., Norman, K.: Improving user satisfaction: The questionnaire for user interaction satisfaction Version 5.5. In: Proceedings of the 1st Annual Mid-Atlantic Human Factors Conference, Virginia Beach, VA, pp. 224–228 (1993)
17. HFRG, Human Factors Research Group. Software Usability Measurement Inventory, SUMI. http://www.ucc.ie/hfrg/questionnaires/sumi
18. Holzinger, A.: Usability engineering for software developers. Commun. ACM, **48**(1), 71–74 (2005)
19. Holzinger, A.: Application of rapid prototyping to the user interface development for a virtual medical campus. IEEE Softw. **21**(1), 92–99 (2004)
20. Hsi, S., Soloway, E.: Learner-centered design: Addressing, finally, the unique needs of learners. In: Computer Human Interaction '98 (CHI '98) conference summary, Los Angeles, USA, pp. 211–212. ACM Press, New York (1998)
21. ISO (1998) Ergonomic requirements for office work with visual display terminals (VDTs)—part 11: Guidance on usability (ISO 9241–11:1998(E)). Geneva, Switzerland
22. Jones, M., Okey, J.: Interface design for computer-based learning environments. http://ddi.cs.uni-potsdam.de/HyFISCH/Multimedia/Learning/InterfaceDesignJones.htm (1995)
23. Kinshuk, P.A., Russell D.: Intelligent and adaptive systems. In: Collis B., Adelsberger H., Pawlowski J. (eds) Handbook on Information Technologies for Education and Training, pp. 79–92. Springer, Berlin (2001)
24. Levi, M., Conrad, F.: Usability testing of World Wide Web sites. Research Papers of Bureau of Labour Statistics. http://stats.bls.gov/ore/htm_papers/st960150.htm (1998)
25. Lewis, J.: IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. Int. J. Hum. Comput. Interact. **7**(1), 57–78 (1995)
26. Macleod, M., Bowden, R., Bevan, N., Curson, I.: The MUSiC performance measurement method. Behav. Inf. Technol. **16**, 279–293 (1997)
27. Murray, T.: Authoring intelligent tutoring systems: An analysis of the state of the art. Int. J. Artif. Intell. Educ. **10**, 98–129 (1999)
28. Nielsen, J.: Jacob Nielsen on e-learning. Elearningpost. http://www.elearningpost.com/features/archives/001015.asp (2001)
29. Nielsen J.: Why you only need to test with 5 users. Alertbox March. http://www.useit.com/alertbox/20000319.html (2000)
30. Nielsen, J.: Usability Engineering. Academic Press, Boston (1993)
31. Notess M.: Usability, User Experience and Learner Experience. eLearn Magazine. ACM Press, New York. http://www.elearnmag.org (2001)
32. Nute, D.: Knowledge representation. In: Shapiro S. (ed): Encyclopedia of Artificial Intelligence, pp 743–869. Wiley, London (1992)
33. Pantazis, C.:(2001) Executive summary: A vision of E-learning for America's workforce. Report of the commission on technology and adult learning, ASTD. http://www.learningcircuits.org/2001/aug2001/pantazis.html (1992)
34. Parlangeli, O., Marchigiani, E., Bagnara, S.: Multimedia systems in distance education: Effects of usability on learning. Interact. Comput. **12**, 37–49 (1999)
35. Patel A, Kinshuk (1996) Knowledge characteristics: Reconsidering the design of intelligent tutoring systems. In: Behrooz, A. (ed.) Knowledge Transfer, pp. 190–197. pAce, London
36. Preece, J., Rogers Y., Sharp H., Benyon D., Holland S., Carey T.: Human–computer Interaction. Addison-Wesley, Wokingham (1994)

37. Quigley, A.: Usability-tested E-learning. Not Until the Market Requires It. eLearn Magazine. ACM Press, New York. http://www.elearnmag.org (2002)
38. Quillian, R.: Semantic memory. In: Minsky M. (eds.) Semantic Information Processing, pp. 216–270. MIT Press, Cambridge (1968)
39. Rickel, J.: Intelligent computer-aided instruction: A survey organized around system components. IEEE Trans. Syst. Man Cybern. **19**(1), 40–57 (1989)
40. Rosić, M.: Establishing of distance education systems within the information infrastructure. M.Sc. Thesis. Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia (in Croatian) (2000)
41. Shackel B.: Usability—context, framework, design and evaluation. In: Shackel B., Richardson S. (eds.) Human Factors for Informatics Usability. Cambridge University Press, Cambridge, 21–38 (1991)
42. Sherwood, R.: Models of computer use in school settings. In: Kinzer, Ch., Sherwood R., Bransford J. (eds.) Computer Strategies for Education: Foundations and Content-area Applications, pp. 105–130. Merrill Publishing Company, Columbus (1986)
43. Shneiderman, B.: Designing the User Interface. Strategies for Effective Human–computer Interaction, 3rd edn. Addison-Wesley, Reading (1998)
44. SIGCHI (2001) Notes from E-learning Special Interest Group (SIG) Discussion at CHI 2001. http://www.elearning.org/
45. Sloan, D., Kelly, B., Heath, A., Petrie, H., Fraser, H., Phipps, L.: Contextual Web accessibility—maximizing the benefit of accessibility guidelines. WWW 2006 Edinburgh, Scotland 22–26 May 2006. In: Conference Proceedings, Special Interest Tracks, Posters and Workshops (CD ROM). http://www.ukoln.ac.uk/web-focus/papers/w4a-2006/ (2006)
46. Squires, D., Preece, J.: Predicting quality in educational software: Evaluating for learning, usability and the synergy between them. Interact. Comput. **11**, 467–483 (1999)
47. Squires, D., Preece, J.: Usability and learning: Evaluating the potential of educational software. Comput. Educ. **27**, 15–22 (1996)
48. Stankov, S.: Isomorphic model of the system as the basis of teaching control principles in the intelligent tutoring system. Ph.D. Dissertation, Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, University of Split (in Croatian) (1997)
49. Stephanidis, C., Salvendy, G., Akoumianakis, D., Bevan, N., Brewer, J., Emiliani, P.L, Galetsas, A., Haataja, S., Iakovidis, I., Jacko, J., Jenkins, P., Karshmer, A., Korn, P., Marcus, A., Murphy, H., Stary, C., Vanderheiden, G., Weber, G., Ziegler, J.: Towards an information society for all: an international R&D agenda. Int. J. Hum. Comput. Interact. **10**(2), 107–134 (1998)
50. Zaharias, P.: E-learning design quality: A holistic conceptual framework. In: Encyclopedia of Distance Learning, vol. II. Idea Group Inc., Hershey (2005)
51. Zaharias, P.: Usability and E-learning. The Road Towards Integration. eLearn Magazine. ACM Press, New York. http://www.elearnmag.org (2004)
52. Zaharias, P., Vassilopoulou, K., Poulymenakou, A.: Designing on-line learning courses: Implications for usability. Sci. J. Appl. Inf. Technol. **1**(1) (2002)
53. WAMMI consortium (2002), WAMMI—web usability questionnaire. http://www.wammi.com
54. Woolf, B.: AI in education. In: Shapiro I., Stuart C. (Eds.) Artificial Intelligence—Encyclopedias, pp. 434–444. Wiley, New York (1992)
55. Wrigh, P., Monk, A.: A cost-effective evaluation method for use by designers. Int. J. Man Mach. Stud. **35**(6), 891–912 (1991)