eLSE Methodology: a Systematic Approach to the Evaluation of e-Learning Systems

Rosa Lanzilotti[1], Carmelo Ardito[1], Maria F. Costabile[1], Antonella De Angeli[2]
[1]Dipartimento di Informatica – Università di Bari – 70125, Bari - Italy
{lanzilotti, ardito, costabile}@di.uniba.it
[2]School of Informatics - University of Manchester Po BOX 88 M601QD - UK
Antonella.de-angeli@manchester.ac.uk

ABSTRACT

Quality of e-learning systems is one of the important topics that the researchers are investigating in the last years. This paper refines the concept of quality of e-learning systems and proposes a new framework, called TICS (Technology, Interaction, Content, Services), which focuses on the most important aspects to be considered when designing or evaluating an e-learning system. Our proposal emphasizes user-system interaction as one of such important aspects. Guidelines that address the TICS aspects and an evaluation methodology, called eLSE (e-Learning Systematic Evaluation) have been derived. eLSE methodology combines a specific inspection technique with user-testing. This inspection, called AT inspection, uses evaluation patterns, called Abstract Tasks (ATs), that precisely describe the activities to be performed during inspection. The results of an empirical validation of the AT inspection technique, carried out to validate this technique, have shown an advantage of the AT inspection over the other two usability evaluation methods, demonstrating that Abstract Tasks are effective and efficient tools to drive evaluators and improve their performance.

Keywords: E-learning systems, quality, evaluation.

INTRODUCTION

E-learning is becoming very important in fields where access to learning materials needs to be brought about effectively and efficiently. Its “any time, any place” nature could be a winning strategy for particular needs, such as decongestion of overcrowded education facilities, support for learners or lecturers who live far from schools and universities, life-long education. When making remote data and tools available to users it is necessary to consider their different characteristics, such as cultural background, technical experience, technological equipment, and physical/cognitive abilities. In the e-learning context, a major challenge for designers and Human-Computer Interaction (HCI) researchers is to develop software tools that can engage novice learners and support their learning even at a distance. Towards this end, there should be a synergy between the learning process and the learner’s interaction with the software. As for any interactive system, usability is a primary requirement. If an e-learning system is not usable, the learner spend more time learning how to use the software rather than learning the contents. Beside being usable, an e-learning system must be effective in meeting the instructor’s pedagogical objectives. System evaluation should thus integrate an assessment of the educational quality aspects of e-learning systems.

Despite the large number of e-learning systems now available, one of the barriers to successful deployment of technology-based learning is the lack of high quality systems tailored to the needs of individual users and groups. Quality, from the Latin *qualis*, which means a pleasant thing, is an abstract term that assumes specific meanings according to the context in which it is used. From the end of the 1970s, in the software engineering context, some factors have been introduced as measures of the software quality. McCall affirms that quality factors represent attributes or characteristics of the software that a user or a client of the software couples with the quality of the software (McCall, 1994). Details on the first studies on quality factors can be found in (McCall, 1994; Boehm, 1978).

When speaking of quality, it is important to consider the regulations for quality certification. In particular, the ISO/IEC 9126 establishes standards for ensuring the quality of a software product (ISO 9126, 1991), emphasizing that the quality is an attribute that depends on the users, the context, the goal, and the cost of the product.

In the last few years, we have been involved in e-learning projects and faced several problems that impact learner-centred design, which presents its peculiarities with respect to general user-centred design (Quintana et al., 2001; Costabile et al., 2003; Ardito et al., 2004; Ardito et al., 2006). The evaluation of e-learning systems deserves special attention, and evaluators need appropriate guidelines as well as effective evaluation methodologies (Zaharias et al., 2002). Unfortunately, the number of studies addressing evaluation of e-learning systems is relatively small and inadequate to the importance of the issue (Squires and Preece, 1999; Quinn et al., 2005). Moreover, it is often the case that the evaluation criteria are only vaguely stated (Parlangeli et al., 1999; Squires and Preece, 1999; Wong et al., 2003), so that an actual measurement of the system quality is left to subjective interpretation.

In this paper, we propose a new framework for quality of e-learning systems, called *TICS* (Technology, Interaction, Content, Services), which focuses on the most important aspects to be considered when designing or evaluating an e-learning system. TICS makes explicit the importance of user-system interaction that is neglected by other authors, who implicitly assume that the user interface of an e-learning system does not influence user activities (Moore, 1989; Ewing and Miller, 2002). Actually, making the user interface 'transparent' is the ambition of any interactive system. In this way, users could concentrate only on their tasks. People who have experience in e-learning know the difficulties of users when interacting with an interface that is not usable. Many e-learning systems provide various functionalities and services, but it is often difficult, if not impossible, for users to find and use them in an effective and efficient way. Thus, our definition of e-learning systems quality highlights the importance of this aspect in designing and/or evaluating the overall quality of these systems. If designers are fully aware of the high value of the user interface, they will not neglect this aspect of e-learning systems and will create products of better quality, which will contribute to the success of technology-based learning.

TICS aims at highlighting the main aspects that contribute to the quality of e-learning systems, so that designers and evaluators can focus on such aspects, thus providing what is required for good quality system. As a further contribution we have derived guidelines that address the TICS aspects and an evaluation methodology, called *eLSE* (e-Learning Systematic Evaluation). eLSE methodology combines a specific inspection technique with user-testing. This inspection aims at allowing inspectors that may not have a wide experience in evaluating e-learning systems to perform accurate evaluations. It is based on the use of evaluation patterns, called Abstract Tasks (ATs), which precisely describe the activities to be performed during inspection. For this reason, it is called AT inspection. An empirical validation of the AT inspection technique has been carried out: three groups of novice inspectors evaluated a commercial e-learning system applying the AT inspection, the heuristic inspection, or user-testing. Results have shown an advantage of the AT inspection over the other two usability evaluation methods, demonstrating that Abstract Tasks are effective and efficient tools to drive evaluators and improve their performance.

The paper has the following organization. First, contributions of some researchers in the domain of e-learning systems evaluation are reported. Then, the TICS framework is illustrated, highlighting the most important aspects to consider in designing and/or evaluating the overall quality of e-learning systems. Consequently, the eLSE methodology is presented and the validation of the AT inspection with its results is reported. Finally, the Section Conclusion closes the thesis.

QUALITY OF E-LEARNING SYSTEMS

Various definitions and frameworks for the quality of e-learning systems are reported in literature, but the identified solutions appear as a short blanket, able to cover only some of the multiple aspects that characterize the complexity of the problem.

Quality is defined as "*the totality of characteristics of an entity that bear on its ability to satisfy stated and implied user needs*" (ISO 9126, 1991). It includes a new quality model distinguishing between three different approaches to product quality:

- *external quality*, which is measured by the dynamic properties of the code when executed (such as response time)
- *internal quality*, which is measured by the static properties of the code, typically by inspection (such as path length)
- *quality in use*, which is measured by the extent to which the software meets the needs of the user in the working environment (such as productivity).

External quality is a result of the combined behaviour of the software and the computer system, while quality in use is the effectiveness, productivity and satisfaction of the user when carrying out representative tasks in a realistic working environment. External measures can be used to validate the internal quality of the software. Quality in use measures the degree of excellence and can be used to validate the extent to which the software meets user needs. Appropriate internal attributes of the software are a pre-requisite for achieving the required external behaviour, and appropriate external behaviour is a pre-requisite for achieving quality in use (ISO 9126, 1991).

Chua and Dyson propose the ISO/IEC 9126 Quality Model to evaluate e-learning systems (Chua and Dyson, 2004; ISO 9126, 1991). It provides an indication to educators and educational administrators of the quality of a system they are considering buying and provides a basis for comparison of different systems. However, for software developers without educational expertise, the ISO model alone would not be sufficient because it is a general software quality model and does not specify the particular teaching and learning activities needed for effective learning. Moreover, ISO/IEC 9126 presents some inherent weaknesses, particularly with regards to the usability characteristic. The authors suggest that this characteristic should be extended to include more specific factors such as consistency, simplicity, legibility. In addition,

they propose the inclusion of user satisfaction as a global characteristic to summarise the general impact of the system on the users in their specific educational context (Chua and Dyson, 2004).
Some authors have highlighted that quality in e-learning needs to be envisaged from different perspectives. However, they look at the quality primarily from a pedagogical point of view and emphasize the learner's perspective (Ehlers, 2004), while they do not analyze in depth aspects related to the design of interactive software systems.
Accordingly, Ehlers (2004) suggests that subjective quality requirements be structured in 7 fields of quality (see Figure 1):

- Quality Field (QF) 1 (Tutor Support) considers the preferences that learners have for communication and cooperation with the tutor of an online course.
- QF 2 (Cooperation and Communication in the Course) contains quality requirements for the course that learners express, that concern the communication and cooperation environment in which they work with other learners in learning groups, with experts or the tutor.
- QF 3 (Technology) is also important to learners. If technical requirements are fulfilled they do not raise the perceived quality very much - as they are taken for granted. Yet if the expected technical standards are not met the learners quality assessment decreases.
- QF 4 (Costs-Expectations-Benefits) refers to the information possibilities learners have about a course or the institution/organization which is offering the course.
- QF 5 (Information Transparency of Provider/Course) contains the provision of formal and standard information as well as individualized counseling on course contents, learning methodology or technical advice.
- QF 6 (Course structure) contains learners' requirements concerning the structure of an e-learning course. Learners' quality preferences clearly show that the presence of lessons as part of an e-learning course (blended learning) is of high importance to certain groups of learners, whereas others do not regard them as important.
- Finally, QF 7 (Didactics) covers aspects of content, learning goals, methods and materials. Experienced e-learners are often very precise in their requirements concerning the didactical setting of an e-learning course.

The quality fields are subdivided into 30 dimensions, each of them represents a set of criteria of learners' preferences that are clustered in a dimension on the basis of empirical evidence.

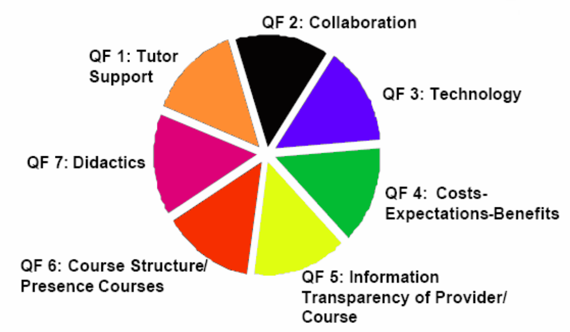

Figure 1. Model of subjective quality requirements from [Ehlers, 2004]

As Ehlers, Rovinskyi and Synytsa have suggested to perform multidimensional quality evaluation according to the parameters identified as important by the learner (Rovinskyi and Synytsa, 2004). These parameters will be hereafter named "quality metadata" and will actually be a part of metadata for a course or derived from the metadata of its components. Quality metadata may be used to measure the correspondence of a course to the needs and expectations of a learner, thus determining the course's quality from a learner's perspective. Quality metadata fall into the following groups: learning goals (learning objectives and tasks, curricula and certificates); instructors and experts (their duties, qualification, and way of interaction); target audience (required knowledge and skills, prior education or professional background, learning groups and interaction); learning environment (required hard and software, additional plug-ins, complexity of environment, navigation, help, assessment instruments and feedback); learning resources (types of resources, availability on/off line, variety of representation, composition and complexity, types of activity offered, learning material completeness).
Other researchers have proposed to consider technological as well as educational aspects in evaluating e-learning systems. In (Herrington et al., 2001) , a framework is presented as a checklist enumerating what are considered to be critical elements of quality learning environments. The checklist is based on the determination of critical elements within three main areas:

- the *pedagogies area*, referring to the activities that are fundamental for the learning process: the learning activities, proposed by the e-learning systems, would involve tasks that reflect the way in which the knowledge will be used in the real life settings; the learners would have the opportunity to work collaboratively; the lecturer's role would be considered as a coach rather than an instructor; and, finally, authentic and integrated assessment tests should be used to evaluate learner' achievement;

– the *resources area*, concerning the content and information provided to the learners: resources should be organized in ways that make them easily accessed. They should be current and based on literature reviews by lecturers and they should present various viewpoints to allow learners the opportunity to judge the merit of different arguments;
– the *delivery strategies area*, involving issues associated with the ways in which the course is delivered to the learners: the learning material should be accessible and available to all learners, including people with special needs, such as people with physical disabilities, or geographically isolated learners. Moreover, the learners should be count on the reliability of the technology: the e-learning system should allow learners to navigate and download materials within a reasonable period of time, and learners should be able to use various technologies to communicate and collaborate with each other and their teachers, and so on.

In a publication that has been approved and publicized by the Italian Ministry of Education, Montedoro and Infante indicate three quality dimensions of e-learning systems: *technology, content and services* (Montedoro and Infante, 2003) Concerning technology, the educational process has to be supported by an on-line platform that manages the educational material. The platform has to be compatible with the international standards (AICC, SCORM, etc.) and to manage functions such as: analysis of the user profiles, building of personalized educational paths, the planning of the learning activities, the management of on-line courses and educational materials, interaction and in particular, forums, chat, virtual classrooms, and communications among learners. Regarding content, it is important to perform an accurate educational design that is learner-centered. It is characterized by the following elements: user orientation, the ease of navigation and fruition of the educational products and resources, interaction in all the phases if the educational process is enriched by specific and continuous feedback, length of the didactic unit not exceeding twenty minutes, content modularity, evaluation of the educational results. Finally, the services dimension increases the learner's possibilities of contacting the tutor. The tutor has the task of managing the relations with the learners and guaranteeing rapid answers.

Grutzner et al. state that the quality of e-learning systems is affected by four dimensions: the *content* of learning materials; the *presentation* of these materials; the *pedagogic content*, i.e., the way in which materials are taught; the overall *functionality* of the courseware. All four dimensions have to be considered at the same time and continuously throughout the courseware life-cycle to ensure high quality of the final product and thus to facilitate learning (Grutzner et al., 2004). The authors propose IntView, a systematic courseware development process, designed to support and integrate all of the different perspectives and views that are involved in the process of designing high quality courseware. The courseware engineering life-cycle model of the IntView methodology is accompanied by four engineering methods that assure the quality of the output artefacts of each phase. These methods include perspective-based inspections, prototyping, tests, and both formative and summative evaluations. Thus, the whole life-cycle is encompassed by verification and validation tasks that are assigned to different roles.

In one of the early papers in distance education, Moore explicitly speaks of user interaction as an important aspect to be considered for the quality of e-learning systems and proposes a quality analysis model that considers three types of interaction (Moore, 1989). First of all learner-content interaction is considered, since it is the process of interacting with content that brings about changes in the learner's understanding and perspective. Learner-lecturer interaction is also important, because lecturers seek to stimulate the learner's interest and to motivate the student to learn. Learner-learner interaction among members of a class is the third type of interaction, since it is often an extremely valuable resource for learning. Like Moore, Fata affirms that for evaluating the quality of e-learning systems it is important to consider three different types for interaction (Fata, 2003): interaction between the learners and didactic material, the educational material just be of good quality, pleasant, easily usable, and equipped with an accurate and rich bibliography; interaction among learners, tutor, and experts: the tutor has to favour discussions among the learners, to reduce the learners' sense of isolation. During the course, the learners must constantly be able to perceive the tutor's presence; and during the conclusion of the course, the tutor must stimulate meta-cognitive reflection; and finally, interaction among learners: the communications among learners should take into account not only the discussions as group activities but also the free communications, in order to encourage a greater familiarity among learners. Other researchers have spoken of interaction as an important aspects for quality e-learning systems (Scanlon et al., 2000; Muirhead, 2001; Ewing and Miller, 2002), but it is worth noticing that they focus on the interaction between learner and educational content.

All these types of interaction are mediated by the software system through which the content is delivered. As HCI researchers, we know the importance of designing the interface to allow a good interaction between the user and the provided content, which is the basis for effective learning. We must suggest models and techniques that support e-learning system designers in creating quality products.

TICS FRAMEWORK FOR THE QUALITY OF E-LEARNING SYSTEMS

With respect to the other definitions of e-learning systems quality, we emphasize the interaction dimension and, specifically, the interaction between the user (teacher or learner) and the overall system, not only its content (the learning materials). On the basis of our experience, we believe that in the e-learning context, design methodologies and techniques have to ensure a user-system interaction that facilitates the learning process. It is not possible to neglect the influence of the user interface on the learner activities. Thus, the interface must be designed and evaluated on the basis of well defined criteria and methods, specific for e-learning (Lanzilotti, 2006).

Our novel definition says that "e-learning systems quality is the extent with which technology, interaction, content and offered services comply with expectations of learners and teachers by allowing them to learn/teach with satisfaction".

By considering the literature on the e-learning systems quality (Briefly reported in the previous section), the experience of human-computer interaction experts and the results of observing real users interacting with e-learning systems, a new framework called *TICS* (*Technology, Interaction, Content, Services*), has been developed. It focuses on the most important aspects to be considered when an e-learning system is designed or evaluated (Figure 2). TICS primarily focuses on the quality dimensions recommended by the Italian Ministry of Education in (Montedoro and Infante, 2003), integrating them with the interaction dimension that, as we have discussed above, plays a crucial role in the fruition of the e-learning material. In the following each TICS aspect is described.

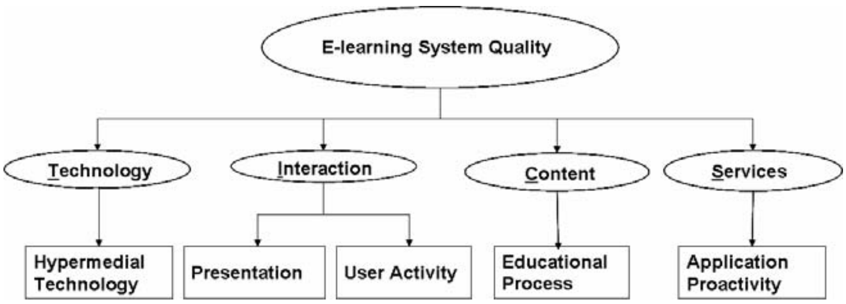

Figure 2. Quality aspects of e-learning systems.

Technology

The technology aspect is specialized in the dimension *hypermedial technology*, that refers to the technological problems that can obstruct or make difficult the use of the e-learning system and to the hypermedia characteristics of e-learning systems i.e., compatibility with different operating systems; performance; monitor resolutions; hardware and software for system accessibility. By technology we also refer to the hypermedia characteristics of e-learning systems, such as selecting and integrating the media made available by current technology. In fact, when designing an e-learning system, we must accurately choose media that are really complementary and may enrich the contents without overloading the user's perceptual and cognitive systems.

Interaction

The explicit inclusion of the interaction aspect is one of the novelty of our approach. We believe it is crucial for technology-based learning. This aspect involves *presentation* of the educational material and of the provided services and tools, and *user activity* performed during the interaction with the system. In particular, user errors should be prevented as much as possible. If an error occurs, the e-learning system should provide appropriate support to manage it. Learners not only interact with their educational material but they consider comments provided by other users. In fact, according to other researchers (Ewing and Miller, 2002; Moore, 1989), a great contribution to learning comes from interaction among learners and their lecturers. Researchers who neglect the importance of user-system interaction implicitly assume that the e-learning system interface does not influence user activities. People who have experience in e-learning know the problems that users have when interacting with an interface that is not usable. In some studies performed with users (Ardito et al., 2006, Lanzilotti, 2006), we have seen that if an e-learning system interface is not usable, learners spend more time learning how to use the software rather than learning the content, which is their main goal.

Content

The content aspect is more closely related to education. It refers to the *educational process* that happens during the interaction with the e-learning system. It focuses on the appropriateness and quality of the educational material that could be achieved through an accurate learner-centered design. This aspect also refers to the way the material is taught and to the capability of the e-learning system to propose study activities to the learner, who should also be free to autonomously choose his/her learning path (e.g. alternating moments of study, exercises, and verifications). Besides, the learner must have the possibility to go more deeply into the course content and the system must provide concrete, real, and helpful examples to facilitate understanding of the educational content.

Services

The services aspect refers to the *application proactivity*, that involves the tools that facilitate and support the user during the navigation through the system and the fruition of the educational material. It is necessary to provide users with communication tools, auto-evaluation tools, help services, search engines, references, scaffolding, and so on. Ease of learn and ease of use of such tools permit users to concentrate their efforts on the learning paths without being required to spend too much time trying to understand the way the e-learning system works.

ELSE METHODOLOGY

Designers and evaluators can focus on TICS dimensions, for providing what is required for good quality system. As a further contribution we have derived guidelines that address the TICS aspects and an evaluation methodology, called eLSE (e-Learning Systematic Evaluation), which prescribes a structured flow of activities. The main idea of eLSE is that the most reliable evaluation results can be achieved by systematically combining inspection with user-based evaluation. Several studies have outlined how two such methods are complementary and can be effectively coupled to obtain a reliable evaluation process (Nielsen, 1993). In line with those studies, eLSE suggests coupling inspection activities and user-testing, and precisely indicates how to combine them to make evaluation more reliable and still cost-effective.
eLSE derives from a usability evaluation methodology, called SUE (Systematic Usability Evaluation), originally developed for evaluating hypermedia systems (Matera et al., 2002). eLSE shares with SUE three important characteristics. First, eLSE couples inspection and user testing, to make the evaluation more reliable and still cost-effective. Each evaluation process starts by having evaluators inspecting the application and identifying possible problems and troubles. User testing is then conducted, whenever necessary, to validate some inspection findings with real users. Second, eLSE suggests to analyze an application along the TICS dimensions (i.e. hypermedial technology, presentation, user activity, educational process, application proactivity) that address the appropriateness of design with respect to the peculiar nature and purposes of the e-learning systems. Finally, eLSE proposes an inspection technique based on the use of Abstract Tasks (ATs) that are specifically defined for e-learning systems. For this reason, it is called *AT inspection*. ATs precisely describe which objects of the application to look for and which actions the evaluator must perform in order to analyze such objects.

ELSE PHASES

According to eLSE methodology, the activities in the evaluation process are organized into a *preparatory phase* and an *execution phase*. The preparatory phase is performed only once for each analysis dimension; its purpose is to create a conceptual framework that will be used to carry out actual evaluations. The output of the preparatory phase can be easily shared among different evaluators, or different evaluation laboratories that have similar interests and evaluate such applications from similar points of view. The preparatory phase consists of the identification of guidelines to be considered for the given dimensions and the definition of a library of ATs. The execution phase is performed every time a specific application must be evaluated. It mainly consists of inspection, performed by evaluators. If needed, inspection can be followed by user testing sessions, involving real users. At the end of each evaluation session, the evaluators must provide designers and developers with an organized evaluation feedback. The activities in the two phases are described in the following sections.

THE PREPARATORY PHASE

In the preparatory phase, a number of decisions must be taken and the definition of a specific set of Abstract Tasks must be carried out.

Abstract Task formulation

By considering TICS dimensions, the literature on e-learning, the experience of human-computer interaction experts, and the results of observing real users interacting with e-learning systems (Ardito et al., 2006), a number of specific guidelines have been identified, to be taken into account during the initial design phase. Then, a set of Abstract Tasks focusing on these guidelines is identified to support the evaluators in their inspections; an AT addresses one or more TICS guidelines. In this way, the defined set of ATs allows inspectors to verify that all the design aspects that must be considered for developing a good e-learning system are taken into account.

An AT is a description of what an evaluator has to do when inspecting an application. ATs guide the evaluator's activities by describing which elements of the application to look for, and which actions the evaluators must perform in order to analyse such elements. In this way, even novice evaluators, with lack of expertise in usability and/or application domain, are able to come out with more complete and precise results. We use the term abstract task since it describes a task that an inspector performs when is evaluating an application, but this description is provided in general terms and abstracts from a specific application to be evaluated.

ATs can be seen as evaluation patterns, making possible to maximize the reuse of evaluators' know-how, by capturing usability inspection expertise, and by expressing it in a precise and understandable form, so that it can be easily reproduced, communicated, and exploited. They therefore allow evaluators to take "... advantage of any of the efforts done in previous works, to reduce the effort needed to achieve a new one" (Nanard et al., 1998).

As stated above, ATs are precisely formulated by means of a template that provides a consistent format and includes the following items:

- *AT Classification Code and Title*: they univocally identify the AT, and succinctly convey its essence.
- *Focus of Action*: it shortly describes the context, or focus, of the AT, by listing the application components that are the evaluation entities.
- *Intent*: it describes the problem addressed by the AT and its rationale, trying to make clear which is the specific goal to be achieved through the AT application.
- *Activity Description*: it describes in detail the activities to be performed during the AT application.
- *Output*: it describes the output of the fragment of the inspection the AT refers to.

Optionally, a comment is provided, with the aim of indicating further ATs to be applied in combination, or when available, significant examples of inspection findings should be reported, to better clarify which situations the evaluators should look for while applying the AT activity.

ATs are defined by expert evaluators on the basis of their experience, complemented with observations of other experts performing various inspections. In this way, ATs are a means for capturing evaluators' expertise and for reusing it; moreover, they provide information about application domain, tasks and users, that is reported in the AT description. Being the heuristic evaluation based only on the use of heuristics, it lacks this information, which is usually provided through some training given to the evaluators prior to the inspection.

As we already said, TICS is the framework to which eLSE also refers. Indeed, the defined ATs are classified in two main categories: *Content learnability* and *Quality in use*. Specifically, content learnability addresses both user activity and educational process. On the other hand, quality in use addresses hypermedial technology, presentation, and application proactivity. The ATs refer to ISO standards (ISO 9126, 1991; ISO 9241, 1997), thus supporting the evaluations of effectiveness, efficiency, security, productivity, and satisfaction.

Table 1 reports an example of AT, whose title is "Graphical interface elements". The AT derives from two guidelines referring to Presentation dimension, namely "Choose text dimension, font, and colour for having good readability" and "Avoid possible forms of distraction (e.g. flashing, sliding inscriptions…)".

QU_02: GRAPHICAL INTERFACE ELEMENTS *Focus of action*: interface graphical elements *Intent*: analyze the platform interface from the graphical viewpoint *Activity description*: Analyze: - the colours - the use of flashing or sliding inscriptions - the characters font and dimension - the coherence of the LO pages. *Output*: a list reporting if: - there is an exaggerated use of different colours - there is an exaggerated use of distraction form (flashing or sliding inscriptions) - characters are not easily readable - the different platform pages are not coherent among them.

Table 1. Example of AT.

Our approach proposes an important distinction between the e-learning platform and the educational modules. The e-learning platform is the software environment that usually offers a number of integrated tools and services for teaching, learning, communicating, and managing learning material. The educational modules, also called Learning Objects, are the specific learning material provided through the platform. Design guidelines and ATs defined for the platform differ from those ones defined for e-learning modules, since different features and criteria need to be considered (Ardito et al., 2006, Lanzilotti, 2006).

Execution phase

Execution phase activities are carried out every time an e-learning system must be evaluated. They include two major jobs: a *systematic inspection* and a *user-based evaluation*.

Systematic inspection is performed by evaluators. During the inspection, the evaluator uses the ATs to perform a rigorous and systematic analysis and produces a report in which the discovered problems are described, as suggested in the AT. The list of ATs provides a systematic guidance to the evaluator on how to inspect an application. Most evaluators are very good in analysing certain features of interactive applications; however, they often neglect some other features, strictly dependent on the specific application category. Exploiting a set of ATs ready for use allows evaluators with limited experience in a particular domain to perform a more accurate evaluation.

User-based evaluation is conducted only when there is disagreement among the evaluators on some inspection findings, so that validation with real users becomes necessary. ATs are still useful since they indicate how to define the Concrete Tasks (CTs for short), i.e. the actual tasks that users are required to perform during the test. A CT is thus simply formulated by considering the *activity description* item of the AT whose application provided contrasting findings; this description is not general as in the AT but it explicitly refers to the application to be evaluated.

Since the AT activity description is a formulisation of the user tasks, starting from this it is immediately possible to formulate experimental tasks which can guide users in the critical situations encountered by the evaluators during inspection. CTs are therefore conceived as a means of actually verifying the impact, upon the users, of the specific points of the application that are supposed to be critical for e-learning quality. In this sense, they make user-based evaluation better focused, so optimizing exploitation of the users resources and helping to obtain a more precise feedback for designers.

During evaluation execution, a sample of users is observed while they are executing CTs and relevant data are collected (users' actions, users' errors, time for executing actions, etc.). The outcome of this is therefore a collection of raw data. In the result summary, these data are coded and organized in a synthetic manner and then analyzed.

The last activity of the execution phase aims at providing the designers and developers of the application with an organised evaluation feedback. The result of this activity is an evaluation report describing the problems detected, possibly revised in the light of the user testing outcome, using the terminology provided in the AT for referring to system objects or interface elements, and for describing critical incidents. This standardised language increases the precision of the report and decreases the risk of misunderstandings.

AT INSPECTION VALIDATION

The advantage of the AT inspection for e-learning systems over other evaluation techniques has been demonstrated by a controlled experiment. The study involved seventy-three senior students of a Human-Computer Interaction (HCI) class at the University of Bari in Italy. They were divided in three groups that were asked to evaluate a commercial e-learning system by applying the AT inspection, or the traditional heuristic evaluation, or a thinking aloud technique. Each group was assigned to one of the three experimental conditions. The heuristic inspection group had to perform an heuristic evaluation exploiting the "learning with software" heuristics (Squires and Preece, 1999). In the user testing group, every evaluator observed a student during the interaction with the e-learning application using the thinking aloud technique. Finally, the AT inspection group used the inspection technique with ATs proposed by eLSE methodology.

In addition, we recruited 25 students of another computer science class, who use the e-learning system, thus acting as users for the user testing.

A week before the experiment, all participants were given a 1-hour demonstration of the application to be evaluated. A few summary indications about the application content and the main functions were introduced, without providing too many details. A couple of days before the experiment, a training session of about one hour introduced participants with the conceptual tools to be used during the experiment. Each group participated in their specific training session.

Data were collected in a group setting, but every participants worked individually. The study consisted of two experimental sessions lasting three hours each. During the first session, participants evaluated the e-learning system applying the technique they were assigned to. Participants of the three groups had to find usability problems in the application, and to record them on a booklet. A day after, each evaluator was asked to type their discovered problems in an electronic form. This was required in order to avoid readability problems during data analysis. In the electronic booklet, each participant reported the description of the problem, where it occurred, and how it was found. Finally, the evaluator gave a severity rating to the problem in the scale from 1 (I do not agree that this is a usability problem at all) to 5 (usability catastrophe).

RESULTS

Two expert usability evaluators independently examined all the electronic forms in order to identify single and unique usability problems. This analysis led to the identification of 247 problems and 49 non problems, or statements which reported not understandable content or unverifiable information. Non-problems accounted for 7% of the statements written by the inspectors applying the heuristic inspection, and 3% of the problems written by the inspectors applying the AT inspection or performing user testing.

On average, participants reported significantly more problems when applying the AT inspection techniques (mean = 21), than when performing user testing (mean = 9.62) or the heuristic inspection (mean = 12.22). The difference is significant as demonstrated by the results of an Anova with evaluation technique as between-subjects factor ($F_{(2,70)} = 25.45$, $p < .001$). Post-hoc tests applying the Bonferroni's correction indicated that the effect was due to the higher number of problems reported in the AT inspection condition, while there was no significant difference between heuristic inspection and user testing.

An overall efficiency index was computed, which reflected the number of problems each participant found in 10 minutes. The Anova indicated that overall efficiency is affected by experimental condition, $F_{(2,70)} = 3.80$, $p<.05$. Post Hoc comparisons (Bonferroni's correction) showed that eLSE was the most efficient technique (mean = 1.19 problems in 10 minutes, standard error = .08). There were no differences between heuristics and user testing, in the order mean = .91, standard error = .08 and mean = .90, standard error = .08.

The study also revealed that different techniques addressed different type of problems. User testing and heuristic inspection helped to highlight problems common to all interactive systems, whereas the AT inspection focused also on specific problems of e-learning. These findings confirm the peculiar characteristics of AT that specifically addresses the e-learning domain, provided that appropriate ATs are defined.

CONCLUSIONS

This paper has discussed the concept of the quality of e-learning systems. In particular, we have refined the e-learning system quality by introducing the aspect of the user-system interaction that had been neglected by some researchers. There is the ambition that the system user interface is "transparent" to the user. Unfortunately, people, that have experience in e-learning, knows difficulties that user meets when s/he interacts with an unusable interface. From this derives the need to emphasize the interaction aspect and the usability of the user interface. Thus, we have defined a framework, called TICS, which focuses on the most important aspects to consider when an e-learning system is designed or evaluated.

The paper have introduced an approach systematic to the evaluation of e-learning systems, called eLSE methodology, that prescribes a structured flow of activities. eLSE proposes an evaluation technique, called AT inspection. Such a technique falls into the category of the methods called "inspection" which do not involve users, but expert evaluators only. This inspection technique uses evaluation patterns, called Abstract Tasks, which make possible to maximize the reuse of the evaluator's expertise and to express it in a precise and understandable form, so that it can be easily reproduced, communicated, and exploited.

The results of the experiment performed to validate the AT inspection have shown that the inspection based on the use of ATs is capable to address specific issues of e-learning better than other techniques such as a specialised heuristic evaluation and user testing.

During the experiment, the students expressed positive opinions on the use of ATs. They said that the guidance offered by ATs allows them to perform the inspection more accurately. Students perceived to have more control on the inspection process and, consequently, they were confident to obtain good results. Furthermore, the two experts that have examined the electronic forms filled by the inspectors noticed that ATs enforce standardization and uniformity of inspection results, since all inspectors are guided by the output item writing their evaluation report.

We already defined a good number of ATs addressing the quality issues of e-learning systems. Other ATs are currently under development that consider the overall user experience in the e-learning domain.

REFERENCES

(Ardito et al., 2004) Ardito, C., De Marsico, M., Lanzilotti, R., Levialdi, S., Roselli, T., Rossano, V., and Tersigli, M. (2004). *Usability of E-Learning Tools*, Proc. of AVI 2004, Gallipoli, Italy, May 25-28, 80-84.

(Ardito et al., 2006) Ardito, C., Costabile, M.F., De Marsico, M., Lanzilotti, R., Levialdi, S., Roselli, T., and Rossano, V. (2006). An Approach to Usability Evaluation of e-Learning Applications, *Universal Access in the Information Society International Journal*, 4 (3), 270 – 283.

(Boehm, 1978) Boehm, B. (1978), *Characteristics of Software Quality*, North Holland Publishing Co., New York, 1978.

(Chua and Dyson, 2004) Chua, B.B., and Dyson, L.E. (2004). *Applying the ISO 9126 Model to the Evaluation of an e-Learning System*, Proc. of ASCILITE 2004, Perth, Australia, December 5-8, 184-190.

(Costabile et al., 2003) Costabile, M.F., De Angeli, A., Roselli, T., Lanzilotti, R., and Plantamura, P. (2003). Evaluating the Educational Impact of a Tutoring Hypermedia for Children, *Information Technology in Childhood Educational Annual*, 289-308.

(Ehlers, 2004) Ehlers, U. (2004). Quality in e-Learning form a Learner's Perspective, *European Journal of Open Distance and E-Learning*, 2004. http://www.eurodl.org/materials/contrib/2004/Online_Master_COPs.html

(Ewing and Miller, 2002) Ewing, J. and Miller, D. (2002). A Framework for Evaluating Computer Supported Collaborative Learning. *Educational Technology & Society*, 5(1), pp. 112-118.

(Ewing and Miller, 2002) Ewing, J., and Miller, D.(2002). A Framework for Evaluating Computer Supported Collaborative Learning, *Educational Technology & Society*, 5(1), 2002, pp. 112-11.

(Fata, 2003) Fata, A. (2003). La Qualita' nella Formazione a Distanza. *Psico-Pratika*, 4. Available at: http://www.humantrainer.com/articoli/anna_fata_qualita_formazione_a_distanza.html.

(Grutzner et al., 2004) Grutzner, I., Weibelzahl, S., Waterson, P. (2004). *Improving Courseware Quality through Life-Cycle Encompassing Quality Assurance*, Proc. of Symposium on Applied Computing (SAC'04), ACM Press, Nicosia, Cyprus, March 14-17, 2004, pp. 946 - 951.

(Herrington et al., 2001) Herrington, A., Herrington, J., Oliver, R., Stoney, S., and Willis, J. (2001). Q*uality Guidelines for Online Courses: The Development of an Instrument to Audit Online Units*, Proc. of ASCILITE 2001, Melbourne, Australia, December 9-12, 2001, pp. 263-270.

(ISO 9126, 1991) International Organization for Standardisation (1991). ISO/IEC: 9126 Information technology - Software Product Evaluation - Quality characteristics and guidelines for their use, 1991.

(ISO 9241, 1997) International Organization for Standardization. (1997). *ISO 9241: Ergonomics Requirements for Office Work with Visual Display Terminal (VDT)* - Parts 1-17.

(Lanzilotti, 2006) Lanzilotti, R., *A Holistic Approach to Designing and Evaluating e-Learning Quality: Usability and Educational Effectiveness*, PhD dissertation, Dip. Informatica, Università di Bari, Bari, Italy, 2006.

(Matera et al., 2002). Matera, M., Costabile, M.F., Garzotto F., and Paolini P. (2002). "SUE Inspection: an Effective Method for Systematic Usability Evaluation of Hypermedia", IEEE Transactions on Systems, Man and Cybernetics - Part A, 32(1), 93-103.

(McCall, 1994) McCall, J.A. (1994). *"Quality factors" - Encyclopaedia of Software Engineering* (John J. Marciniak, Ed), John Wiley , New York, 958-969.

(Montedoro and Infante, 2003) Montedoro, C. and Infante, V. (2003). Linee Guida per la Valutazione di Qualità del Software didattico nell'e-Learning. ISFOL, I libri del Fondo Sociale Europeo, Roma.

(Moore, 1989) Moore, M.G.(1989). Three Types of Interaction, *The Amer. Journal of Distance Education*, 3 (2), 1989, pp. 1–6.

(Muirhead, 2001) Muirhead, B. (2001). Interactivity Research Studies. *Educational Technology & Society*, 4(3), pp. 108-112.

(Nanard et al., 1998) Nanard, M., Nanard, J., and Kahn, P.(1998). P*ushing Reuse in Hypermedia Design: Golden Rules, Design Patterns and Constructive Templates*, Proc. of Hypertext '98: Proceedings of the ninth ACM conference on Hypertext and hypermedia: links, objects, time and space, Pittsburgh, PA, USA, June 20-24, 1998, 11-20.

(Nielsen, 1993) Nielsen, J. (1993). Usability Engineering, Academic Press, Cambridge, MA.

(Parlangeli et al., 1999) Parlangeli, O., Marchigiani, E., and Bagnara, S. (1999). Multimedia System in Distance Education: Effects on Usability. *Interacting with Computers*, Elsevier Science Ltd, Great Britain, 12, 37-49.

(Quinn et al., 2005) Quinn, C.N., Alem, L., and Eklund, J. (2005). A pragmatic evaluation methodology for an assessment of learning effectiveness in instructional systems. http://www.testingcentre.com/jeklund/Interact.htm.

(Quintana et al., 2001) Quintana, C., Carra, A., Krajcik, J., and Elliot, S. (2001) Learner-Centred Design: Reflections and New Directions, *Human-Computer Interaction in the New Millennium*. In Carroll (Ed.) ACM Press, New York: Addison-

Wesley, 605-626.

(Rovinskyi and Synytsya, 2004) Rovinskyi, D., and Synytsya, K. (2004). *Distance Courses Quality: A Learner's View*, Proc. of Fourth IEEE ICALT'04, Joensuu, Finland, August 30-September 1, 2004, pp. 1080-1081.

(Scanlon et al., 2000) Scalon, E., Jones, A., Barnard, J., Thompson, J., Calder, J. (2000). Evaluating Information and Communication Technologies for Learning. *Educational Technology & Society*, 3(4), 101-107.

(Squires and Preece, 1999) Squires, D. and Preece, J. (1999). Predicting quality in Educational Software: Evaluating for Learning, Usability, and the Synergy between them. *Interacting with Computers*, Elsevier Science Ltd, Great Britain, 11 (5), 467-483.

(Wong et al., 2003) Wong, B., Nguyen, T.T., Chang, E., and Jayaratna, N. (2203). *Usability Metrics for E-Learning*. Workshop on Human Computer Interface for Semantic Web and Web Applications, Springer-Verlag, Heidelberg, Germany, LNCS No: 2889, 2003, 235 – 252.

(Zaharias et al., 2002) Zaharias P., Vasslopoulou K., and Poulymenakou A. (2002). *Designing On-Line Learning Courses: Implications for Usability*. www.japit.org/zaharias_etal02.pdf, 2002.