A COMPARISON OF USABILITY EVALUATIONMETHODS FOR E-LEARNING SYSTEMS

PLANTAK VUKOVAC, D.; KIRINIC, V. & KLICEK, B.

Abstract: *Usability evaluation of e-learning systems has specific requirements that differentiate it from evaluation of other interactive systems. In situations when teachers want to evaluate the usability of their own e-learning courses, it is therefore not easy for them to choose the appropriate evaluation method.*
Lately, severalusabilityevaluation methods adapted for the context of e-learning have been proposed. This paper examines their characteristics and identifies the criteria for choosing the most appropriate methods. While comparing the current usability evaluation methods for e-learning it was established that a lot of methods do not address all the specific issues relevant for e-learning systems and educational modules. Moreover, many methods do not provide sufficient information about the practical application of the method that could be useful to usability practitioners.

Key words: *usability evaluation, e-learning, evaluation criteria*

Authors´ data: M.Sc. **Plantak Vukovac**, D[ijana]; Dr. Ph.D. **Kirinic**, V[alentina]; Univ.Prof. Ph.D. **Klicek**, B[ozidar]; University of Zagreb, Faculty of Organization and Informatics, Pavlinska 2, HR-42000, Varazdin, Croatia, dijana.plantak@foi.hr, valentina.kirinic@foi.hr, bozidar.klicek@foi.hr

This Publication has to be referred as: Plantak Vukovac, D[ijana]; Kirinic, V[alentina] & Klicek, B[ozidar] (2010). A Comparison of Usability Evaluation Methods for e-Learning Systems, Chapter 27 in DAAAM International Scientific Book 2010, pp. 271-288, B. Katalinic (Ed.), Published by DAAAM International, ISBN 978-3-901509-74-2, ISSN 1726-9687, Vienna, Austria
DOI: 10.2507/daaam.scibook.2010.27

1. Introduction

The main research question in the field of human-computer interaction (HCI) is "how to work with and improve the usability of interactive systems" (Hornbæk, 2006). According to the ISO 9241-11standard, usability is defined as the "extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (ISO 9241-11, 1998).

With the divergence of Web-based systems, the focus of HCI research has shifted towards Web usability and the development of methods for usability evaluation of Websites, aimed at specific Web domains, such as e-commerce, tourist, cultural heritage Websites etc.

As the Web has also become a new learning environment, different issues have arisen to be considered in order to fully exploit its advantages and enhance the quality of learning and teaching. One of the aspects that is neglected when evaluating the overall quality of e-learning courses is e-learning usability, resulting in relatively scarce researches examining the usability issues of e-learning applications (Granić, 2008; Kukulska-Hulme & Shield, 2004; Zaharias, 2006). Since the purpose of e-learning systems is not only to interact, but also to support knowledge dissemination and acquisition, traditional usability design guidelines and usability evaluation methods (UEMs) established in the HCI field are not sufficient in the e-learning context (Granić, 2008; Hornbæk, 2006; Zaharias, 2006).

In recent years, researchers have made efforts to develop new sets of guidelines and UEMs suitable for the e-learning domain, considering specific requirements such as the learning process, instructional design, motivation, pedagogical issues etc. However, it seems that a "consolidated evaluation methodology of e-learning applications does not yet exist" (Ardito et al., 2006), and that current method proposals lack a comprehensive and systematic approach that evaluates different e-learning perspectives, which makes it difficult to select useful tools for quick and reliable usability evaluation.

As its aim is to provide a broader view of the aforementioned problems, this paper examines e-learning usability evaluation methods that have emerged lately and proposes a set of criteria that should be consulted when choosing the appropriate method for usability evaluation of e-learning systems or when developing comprehensive new e-learning UEMs.

1.1 Traditional Usability Evaluation Methods

Usability evaluation methods are used for identifying usability problems and improving the usability of an interface design. In general, methods are categorized as analytical or empirical. Analytical methods, also known as inspection methods, are used for interface inspection by usability experts, and are perceived as a quick and low-cost alternative to empirical methods, where testing with actual users is performed.

In continuation, two methods from the inspection methods category and two user testing methods, all of which are found in usability studies of e-learning systems, are briefly described:

- *Heuristic evaluation (HE)* –an informal, cheap and quick method where a small group of usability evaluators inspect a user interface to find and rate the severity of usability problems using a set of usability principles or heuristics (Holzinger, 2005; Nielsen, 1994). It enables the identification of major and minor problems and can be used early in the development process. Its disadvantages are that evaluators have to be experts to provide good results (Hollingsed & Novick, 2007; Holzinger, 2005) and identification of domain-specific problems is not reliable (Holzinger, 2005).
- *Cognitive walkthrough*–enables evaluators' analysis of a user interface by means of simulating step-by-step user behavior for a given task. The emphasis is on cognitive issues,through analyzing the user's thought process (Holzinger, 2005). Its drawback is that it does not provide guidelines and evaluation is not effective if scenarios are not adequately described (Hollingsed & Novick, 2007).
- *Thinking-aloud* –used in usability testing with actual users, where they verbalize their thoughts while interacting with the interface. It enables evaluators to understand how users view the system and why they do something. The method is time-consuming and to some extent unnatural when used, but that disadvantage is neutralized by co-discovery learning where two users use and comment the interface together (Holzinger, 2005).
- *Questionnaire* –used when the subjective satisfaction of users with the interface is measured. Rather than evaluatingthe user interface, it evaluates users' opinions, preferences and satisfaction. Results can be statistically measured, but a large number of responses have to be collected in order to ensure significance (Holzinger, 2005).

There is a consensus that heuristic evaluation identifies more interface problems, and does it more cheaply and sooner than empirical testing, which identifies more severe issues that will likely hinder the user, but at a higher cost (Hollingsed & Novick, 2007). Also, testing users' interaction with an interface should have precedence over users' opinions of what they think they do (Holzinger, 2005), although, on the other hand, questionnaires provide feedback about user satisfaction, which has been found to be a significant factor in students' decision to drop out from e-learning courses(Levy, 2007).

Many authors agree (Granić, 2008; Hollingsed & Novick, 2007; Holzinger, 2005; Ssemugabi & de Villers, 2009; Triacca et al., 2004) that usability inspection should be accompanied by user testing for more reliable results. However, when only one method has to be selected, cost-effective and easy to conduct heuristic evaluation seems to have an advantage (Hollingsed & Novick, 2007; Ssemugabi & de Villers, 2009).

1.2 Specifics of usability evaluation of e-learning systems

According to the ISO definition of usability, three usability constructs can be distinguished: the context of use, the user, and her/his goals. In the context of e-learning, the role of the user is many fold: learner, course designer/teacher, or an e-learning platform administrator. The user has different goals in every role she/he plays: to learn and to test the knowledge, to implement educational content, or to administer e-learning platform and e-courses. The e-learning context is also heterogeneous: it refers to different tools used to accomplish the goals (e.g. e-learning platforms like course Websites, intelligent tutoring systems, learning management systems(LMSs), or educational applications on a CD) and social and physical environment (blended learning environment, online and mobile learning environment).

In order to properly address all important usability issues in different users' roles, particularly the first two (learner and teacher/designer), current design guidelines and usability evaluation methods should integrate cognitions from other fields such as pedagogy, psychology, education, multimedia learning etc. Among the first researchers who noticed that the existing web heuristics could not simply apply to e-learning context were Squires and Preece (1999). They proposed a set of 'learning with software'heuristics by adapting Nielsen's ten heuristics to socio-constructivist criteria for learning. The adaptation of Nielsen's heuristics and heuristic evaluation method to the e-learning context can also be found in other studies (Albion, 1999; Reeves et al., 2002,Ssemugabi& de Villers, 2009).

Other researchers based their evaluations on usability testing, such as (Nokelainen, 2006; Zaharias, 2006), exploring users' perception, satisfaction and motivation to learn. These authors emphasize another usability aspect important in the context of e-learning, the so-called pedagogical usability. While general or technical usability is concerned with usability of virtual environments, i.e the user interface of thee-learning platform, pedagogical usability is concerned with "whether the tools, content, interface and the tasks of the web-based learning environments support various learners to learn in various learning context according to selected pedagogical objectives" (Silius et al., 2003).The main assumption that lies beneath pedagogical usability is "how the functions of the system facilitate the learning of the material it is delivering" (Nokelainen, 2006).Ardito et al. (2006) emphasize that evaluatingthe usability of an e-learning application includes taking into account the e-learning platform and educational content provided through it, while Nokeilainen (2006) claims that the latter is much less frequently studied.

Several researchers acknowledged the benefits of combining usability inspection with user testing, employing two or more evaluation methods in their e-learning usability studies (Ardito et al., 2006; Granić, 2008; Lanzilotti et al., 2005; Ssemugabi & de Villers, 2009; Triacca et al., 2004). However, the e-learning usability area is still maturing (Ssemugabi & de Villers, 2009) and the selection of appropriate UEMsand measures for e-learning courses presents a challenging, if not a difficult task.

2. Research methodology

According to Kothari (1990, p. 8) "research methodology is a way to systematically solve the research problem" and "research methods do constitute a part of the research methodology". "Research methods may be understood as all those methods/techniques that are used for conduction of research" (ibid) and can be put into three groups: 1) methods for data collection; 2) statistical techniques for establishing relationships between the data and the unknown, and 3) methods for evaluating the accuracy of the results obtained.

For the researcher it is important to know which methods and techniques are valuable for the research, to "understand the assumptions underlying various techniques" and to "know criteria by which they can decide that certain techniques and procedures will be applicable to certain problems and others will not" (ibid).

This research was conducted to:

- identify and structure criteria that should be taken into account when choosing the most appropriate methods/methodologies for usability evaluation of e-learning systems,
- identify and analyze existing usability evaluation methods/methodologies for e-learning systems,
- compare e-learning UEMs according to identified criteria.

It must be emphasized that not all authors that proposed UEMs for e-learning use the term methodology to describe different aspects and procedures used to evaluate e-learning systems usability. Some authors use the term method. In other HCI researches, usability evaluation methods are sometimes called techniques. Thus, in order to be consistent in terminology and avoid unnecessary listing, we used the general term 'methods' to refer to all methods, methodologies or frameworks addressed in this paper regardless of their scope. The paper does not address the quality factors of a particular method.

In order to structure the criteria relevant for the selection of the method for usability evaluation of e-learning systems, three starting points were chosen – key criteria and questions used in (Holzinger, 2005), (Dix et al., 2004) and (Preece et al., 2002).

Holzinger (2005) provided the following criteria for comparison of usability evaluation techniques: 1) Applicably in Phase, 2) Required Time, 3) Needed Users, 4) Required Evaluators, 5) Required Equipment, 6) Required Expertise, 7) Intrusive. The descriptions of criteria used for comparison of the methods were not provided.

Dix et al. (2004, pp. 357-360) described these criteria:

1) Stage in the cycle at which the evaluation is carried out (design vs. implementation) – refers to evaluation throughout the design process; ensuring early evaluation brings the greatest pay-off since problems can be easily resolved at this stage;
2) Style of evaluation (laboratory vs. field studies) – refers to decision between controlled experimentation in laboratory and field study, or including both;

3) Level of subjectivity or objectivity of the technique (subjective vs. objective) – considers knowledge and expertise of the evaluator; recognizing and avoiding evaluator bias;
4) Type of measures provided (qualitative vs. quantitative measures) – relates to "the subjectivity or objectivity of the technique, with subjective techniques tending to provide qualitative measures and objective techniques, quantitative measures";
5) Information provided – considers the level of information or feedback required from an evaluator (e.g. low-level, high-level);
6) Immediacy of the response – refers to methods of recording "the user's behavior at the time of the interaction itself", e.g. think aloud method, and methods relying "on the user's recollection of events", e.g. a post-task walkthrough (ibid, p. 359);
7) Level of interference implied (intrusiveness) – the intrusiveness of the technique itself relates to the immediacy of the response;
8) Resources required – respecting the availability of resources: equipment, time, money, participants, expertise of the evaluator and context.

Preece et al. (2002, p. 350) proposed the following criteria when choosing the evaluation paradigm and techniques:

1) Users – refers to involvement of *appropriate* users ensuring that they represent the targeted user population (difference in experience, sex, age, culture, education, personality); determining how the users will be involved in evaluation (place, duration);
2) Facilities and equipment – includes the equipment used in evaluation (e.g.video camera, logging software, questionnaire forms etc.);
3) Schedule and budget constraints – refers to planning evaluations that can be completed on time and within the budget;
4) Expertise – considers the level of expertise of the evaluation team.

"Usability of e-learning poses its own requirements, hence its usability evaluation is different from that of general task-oriented systems and requires different criteria" (Ssemugabi & de Villers, 2009). Thus, another set of criteria relevant in the context of e-learning were compiled by the authors of this paper, based on an extensive literature review and analysis of several proposed e-learning usability evaluation methods. These criteria are:

1) Method instrument(s) – HCI methods and techniques used in the usability evaluation method for e-learning systems;
2) Formal method background – references to other methods, standards, frameworks that enabled the creation of the method constructs;
3) Heuristics/guidelines for evaluation – refers to the development of a set of usability criteria, heuristics or guidelines that enable evaluation;
4) Pedagogical criteria integration – inclusion of pedagogical criteria in evaluation (e.g. learning outcomes, learner control, collaborative learning, motivation, assessment, feedback);
5) Evaluation target – the subject of evaluation (e.g. e-learning platform, e-learning content, or both);

6) Evaluation of stakeholders' roles – evaluator's and/or user's profiles (e.g. usability expert, teacher, course designer, administrator, end-user);
7) Empirical evidence of the method – whether the method has been empirically tested on the actual e-learning system;
8) Empirical comparison to other methods– whether the method has been compared against other UEMs to confirm that the procedure employed for a specific test is suitable for its intended use and/or provides better problem identification;
9) Future developments of the method – indicates the presence and plan of new empirical evidence or method validations.

All four sets of criteria for UEM selection were compared to find analogous criteria. The final set of criteria (divided into general and specific criteria) is presented in Table 1.

Next, several e-learning UEMs were selected for further analysis and comparisons. To enter that procedure, a method had to satisfy the following criteria:

- it extends the existing or proposes a new set of usability tools adapted to the context of e-learning. Studies that evaluated e-learning applications without further modification of traditional HCI methods to the requirements of e-learning domain (e.g. new evaluation criteria, extended guidelines, pedagogical perspective, different user roles) were excluded from the analysis, and
- it is empirically tested in Web-based learning environments.

According to the above criteria and an extensive review of relevant journals and proceedings, the following methods, methodologies or approaches were identified and selected: *SUE methodology* (Ardito et al., 2004; Ardito et al., 2006; Costabile et al., 2005), *eLSE methodology* (Lanzilotti et al., 2006), *MiLE+ method* (Bolchini & Garzotto, 2008; Inversini et al., 2006; Triacca et al., 2004), "Multi-faceted framework for usability evaluation of e-learning applications" (Ssemugabi & de Villers, 2009), "Usability evaluation method for e-learning applications" (Zaharias, 2006) and "PMQL – Pedagogically Meaningful Learning Questionnaire" (Nokelainen, 2006).

However, since we did not systematically check all major HCI-related journals and proceedings of major HCI-related conferences, we might have also missed some relevant method. This can be regarded as one limitation of the study. Another limitation is the absence of criteria related to quality factors of the method, e.g. thoroughness, effectiveness etc., which is an important issue thatneeds to be examined in its own right. Nevertheless, we believe that this paper presents a solid overview of characteristics of the existing e-learning UEMs proposed lately.

Criteria		**Identified in**
GENERAL CRITERIA	Applicably in phase	Dix et al. (2004), Holzinger (2005)
	Required time	Dix et al. (2004), Holzinger (2005), Preece et al. (2002)
	Required budget	Dix et al. (2004), Preece et al. (2002)
	Needed users	Dix et al. (2004), Holzinger (2005), Preece et al. (2002)
	Requirede valuators	Holzinger (2005)
	Required expertise	Dix et al. (2004), Holzinger (2005), Preece et al. (2002)
	Required equipment	Dix et al. (2004), Holzinger (2005), Preece et al. (2002)
	Method intrusiveness	Dix et al. (2004), Holzinger (2005)
	Style of evaluation	Dix et al. (2004), Preece et al. (2002)
	Level of subjectivity or objectivity of the technique	Dix et al. (2004)
	Type of measures provided	Dix et al. (2004)
	Information provided	Dix et al. (2004)
	Immediacy of the response	Dix et al. (2004)
E-LEARNING SPECIFIC CRITERIA	*Method instrument(s)*	Ourcriteria
	Formal method background	
	Heuristics/guidelines for evaluation	
	Pedagogical criteria integration	
	Evaluation target	
	Evaluation of stakeholders' roles	
	Empirical evidence of the method	
	Empirical comparison with other methods	
	Future developments of the method	

Tab. 1. Criteria forcomparison of usability evaluation methods for e-learning systems and modules

3. Review of selected methods for usability evaluation of e-learning systems

3.1 SUE Methodology

Systematic Usability Evaluation or the SUE methodology, primarily developed for usability evaluation of hypermedia systems, combines inspection with user-based evaluation (Ardito et al., 2004; Ardito et al., 2006; Costabile et al., 2005). Usability evaluation is performed in two phases: a *preparatory phase* and an *execution phase*. In the preparatory phase a conceptual framework for evaluation is created by identifying usability attributes for analysis dimensions, considering the application's domain. For each dimension, general usability principles (effectiveness and

efficiency) are decomposed into finer-grained criteria, where a number of usability attributes or guidelines are associated to these criteria. Evaluation is performed using evaluation patterns, called *Abstract Tasks* (AT), addressing the identified guidelines (Ardito et al., 2004). This approach supports the evaluator in analyzing specific components of the application and enables comparison of identified problems performed by different evaluators. The preparatory phase is performed only once for a specific application. The execution phase is performed every time the application is evaluated, consisting of usability inspection, as an obligatory part of the phase, and user testing, that may occur only in critical cases. Inspection is driven by ATs and each inspector should prepare a report describing usability problems.

By adapting the SUE methodology for evaluation of e-learning applications, four analysis dimensions are identified (Ardito et al., 2004; Ardito et al., 2006): *presentation, hypermediality, application proactivity* and *user activity.* For each e-learning dimension general usability principles, criteria and guidelines are derived. ATs enable to evaluate features of an e-learning platform and e-learning modules and are grouped in three categories: *content insertion and content access, scaffolding,* and *learning window*.

Papers describing the SUE methodology (Ardito et al., 2004; Ardito et al., 2006; Costabile et al., 2005) give many details about the methodology's constructs, but no details regarding resources, facilities and equipment, etc., needed to perform evaluation of e-learning applications.

3.2 eLSE Methodology

Another systematic approach for e-learning systems evaluation is eLSE methodology (*e-Learning Systematic Evaluation*) (Lanzilotti et al., 2006). The eLSE methodology is derived from the SUE methodology and shares three characteristics with it: 1) usability inspection, which is the central point of evaluation, followed by user testing; 2) e-learning dimensions, called TICS, which describe almost the same concepts as SUE, but under different names: *Technology* (hypermediality in SUE), *Interaction* (combines presentation and user activity in SUE), *Content* (refers to educational process, partially covered by hypermediality in SUE) and *Services* (application proactivity in SUE); 3) inspection guided by ATs that address one or more TICS guidelines. Like SUE, this methodology enables the evaluation of an e-learning platform and educational modules (learning objects).

Usability evaluation process is also organized in the preparatory phase, where ATs are defined, and the execution phase. In the execution phase, a systematic inspection is performed using ATs classified in two categories: *Content learnability* and *Quality in use*. After inspection, user testing is performed when disagreement about identified problems occurs between inspectors. Users are observed while performing *Concrete Tasks* (CT) formulated from identified critical ATs by inspectors. The evaluation process is finished when the evaluation report is generated describing usability problems detected in AT inspection and possibly during user testing.

The AT inspection part of the eLSE methodology has been validated against heuristic evaluation and thinking aloud, resulting in more usability problems

identified by AT inspection, discovering problems specific to the e-learning domain as well (Lanzilotti et al., 2006). However, no evidence of further methodology development has been identified.

3.3 MiLE+ Method

MiLE+ method (acronym for *Milano-Lugano Evaluation* method) is another method that integrates techniques and evaluation strategies from various 'traditional' usability evaluation methods (heuristic evaluation, scenario driven evaluation, cognitive walkthrough, and task based testing) (Bolchini & Garzotto, 2008; Inversini et al., 2006; Triacca et al., 2004), similarly as SUE and eLSE. Originally, the MiLE method was developed for Web application evaluation, evolving to MiLE+ on the basis of MiLE and SUE methodology concepts. The method has been adapted and applied to usability evaluation of e-learning Web-based systems as well (Inversini et al., 2006; Triacca et al., 2004).

The focus of MiLE+ is on usability inspection performed through two usability activities (Bolchini & Garzotto, 2008): 1) *requirements-independent analysis* or *Technical Inspection*, where usability is evaluated from a 'technical' and 'objective' point of view, and 2) *requirements-dependent analysis* or *User Experience Inspection*, where usability is examined in terms of fulfillment of specific needs of specific users in specified contexts of use.

Technical Inspection applies random heuristic evaluation or, preferably, scenario-based evaluation using technical heuristics organized into six design dimensions (navigation, content, technology, semiotics, cognitives and graphics), which share similarities with SUE and eLSE dimensions. After that, the inspector performs *User Experience Inspection* to 'put him/herself into the shoes' of the user to anticipate problems encountered by end-users during their experience with an application.

Main constructs of MiLE+ are as follows: 1) *scenarios* – they are defined on macro and micro levels to identify user types, profiles and their goals within the context of use. In the context of e-learning, user types are learner and instructor/teacher. The result of scenario definition is a structured set of tasks and goals associated to each user profile;2) *heuristics* – scenarios are supported by heuristics (usability guidelines/principles) that guide technical inspection and user experience inspection; 3) *Usability Evaluation Kits* (U-KITs) –a library of specific evaluation tools comprised of a library of *Technical Heuristics* with 82 heuristics, a library of *User Experience Indicators* with 20 indicators, and a library of scenarios (User Profiles, Goals and Tasks) related to a specific domain.

After inspection, user testing may be performed for the most critical scenarios, goals and tasks identified by inspectors.

The MiLE+ method is more systematic and structured than other evaluation techniques (Bolchini & Garzotto, 2008) and is being constantly revised. Its advantage is in the reuse of scenarios and heuristics, which makes it suitable for novice evaluators. In the context of e-learning, it supports inspection with different user roles, but does not offer many heuristics for the pedagogical perspective or instructional design.

3.4 Multi-faceted framework for usability evaluation of e-learning applications

The framework for usability evaluation proposed in (Ssemugabi & de Villers, 2009) also combines several usability evaluation methods, applying a different approach thanthat in the methods above. In previous methods user testing comes after heuristic evaluation and is performed only for critical problems, while here user testing was performed by applying a questionnaire independently of heuristic evaluation results. The goal of the study was not to create a new e-learning UEM, but compare the results of different usability evaluation methods adapted to the context of e-learning.

The framework for evaluation was the same for evaluators and users and was based on a set of 20 usability criteria defined in three categories: 1) 'learning with software' heuristics from (Squires &Preece, 1999), 2) Website-specific criteria for educational Websites, and 3) learner-centered instructional design criteria. For each criterion a list of sub criteria or guidelines were generated.

First, user testing with students using an e-learning application during the semester was performed, applying a questionnaire with a 5-point Likert scale to measure learners' perception of usability according to criteria from the defined framework. After that, a focus group interview with 8 students was performed to clarify the problems identified.

Also, the heuristic evaluation took place with two evaluators having expertise in HCI and two 'double experts' with expertise in HCI, instructional design and teaching. The experts evaluated an e-learning application independently after familiarizing with the set of heuristics (excluding the criteria regarding personal learning experience), the evaluation process and the e-learning application. After the results of evaluators were aggregated and compiled with the learners' problems identified by the questionnaire, a final list of usability problems was given to evaluators to rate the severity of problems.

The analysis of results showed that 4 evaluators identified more usability problems (77% out of total 75 problems) than 61 learners (73% out of total 75 problems), which gives the heuristic evaluation method advantage in terms of effectiveness, efficacy and cost (Ssemugabi & de Villers, 2009).

This evaluation approach could be enhanced by conducting user testing during the users' actual interaction with the interface, since the assessment of the interface based on memory recall is not reliable (Holzinger, 2005).

3.5 Usability evaluation method for e-learning applications

This generically named method has been developed for user testing considering two aspects: cognitive and affective. The method is actually a psychometric-type questionnaire used to measure learners' perception of e-learning applications usability and learners' intrinsic motivation to learn (Zaharias, 2006). The questionnaire has been designed on the postulates of the *ARCS Model of Motivational Design* by Keller and *Questionnaire Design Methodology* by Kirakowski and Corbett (Zaharias, 2006; Zaharias, 2009).

The main questionnaire's constructs were extracted from a conceptual framework, which employs the following parameters: *Navigation, Learnability,*

Accessibility, Consistency, Visual Design, Interactivity, Content & Resources, Media Use, Learning Strategies Design, Instructional Feedback, Instructional Assessment and *Learner Guidance & Support*. The questionnaire has undergone several versions, with the last one containing 39 items measuring e-learning usability parameters (Web design and instructional design parameters) and 10 items measuring motivation to learn. The method has been validated in two pilot studies in corporate settings on asynchronous e-learning applications. So far, it has not been established whether the method was combined with other usability evaluation methods.

3.6 PMQL – Pedagogically Meaningful Learning Questionnaire

The questionnaire named *Pedagogically Meaningful Learning Questionnaire* (PMQL) is a method aimed at measuring subjective user satisfaction with e-learning platform and e-learning materials (Nokelainen, 2006). It was developed on the basis of usability criteria addressing technical usability and pedagogical usability. The focus is on the assessment of pedagogical usability of digital learning materials through ten dimensions: *Learner control, Learner activity, Cooperative/ Collaborative learning, Goal orientation, Applicability, Added value, Motivation, Valuation of previous knowledge, Flexibility* and *Feedback.*

The questionnaire has undergone two instances of empirical psychometric testing and its final version contains 56 items that measure user satisfaction on a five-point Likert scale. Items that measure issues about e-learning system or issues about content are clearly distinguished. Empirical testing of the questionnaire was performed in elementary school settings.

The criteria developed for PMQL can also be used as the basis for heuristic evaluation (ibid). However, no evidence of further research in that direction or new questionnaire revisions has been found.

4. A comparison of usability evaluation methods for e-learning systems

The methods described in the previous section were compared against the general and specific criteria crucial for choosing the most appropriate method proposed in Section 2.The results of comparison are presented in Table 2. The data about the method regarding particular criteria not found in the available papers are marked with N/A. Some data were not explicitly described in the papers but could be extracted as tacit knowledge. Those data are marked with an asterisk (*).

As seen in Table 2, for the majority of e-learning UEMs general data about the practical applicability of the method, e.g. different resources required, are not available or can only be assumed. Also, some of the methods do not address all specific issues relevant for e-learning systems and modules. In the context of e-learning, methods should enable identification of usability problems related not only to useri nterface, but to learning and pedagogy as well. Different roles that users can have should also be addressed and that aspect is present only in the MiLE+ method. Furthermore, the majority of proposed e-learning UEMs are not compared agains t o ther UEMs to justify their advantages.

Criteria	e-Learning usability evaluation methods					
	SUE methodology	**eLSE methodology**	**MiLE+ method**	**Ssemugabi & de Villers´s framework**	**Zaharias´s Questionaire**	**Nokelainen´s PMQL**
Method instrument(s)	Scenario-based inspection by evaluators using evaluation patterns (ATs) + user testing (optional)	Scenario-based inspection by evaluators using evaluation patterns (ATs) + user testing (optional)	Scenario-based inspection by evaluators using *Usability Evaluation Kit*+ user testing (optional)	Adapted HE inspection by evaluators + user testing (questionnaire)	Psychometric-type questionnaire to measure users' perception of e-learning usability and related affect	Psychometrically tested Pedagogically Meaningful Learning Questionnaire
Formal method background	HE + personal experiences of authors	SUE methodology and TICS framework recommended by the Italian Ministry of Education	SUE and eLSE methodology	Criteria for e-learning usability evaluation from different sources	ARCS Model of Motivational Design and Questionnaire Design Methodology + criteria for evaluation from different sources	Criteria for e-learning usability evaluation from different sources
Heuristics/ Guidelines forevaluation	Yes, analysis dimensions → principles → criteria → guidelines + definition of ATs	Yes, TICS framework → usability dimensions → (not further specified) + definition of ATs/CTs	Yes, library of *Technical Heuristics, User Experience Indicators* and library of scenarios	Yes, conceptual framework (general interface design, Website-specific design, and instructional design) containing 20 criteria: 15 for adapted HE and 20 for the questionnaire	Yes, conceptual framework (Web design and instructional design parameters, and motivation to learn construct) as a basis for the questionnaire containing 49 items (v3)	Yes, 51 sub dimensions as a basis for the questionnaire containing 56 items (v2)
Pedagogica lcriteria integration	Partially	Partially	Partially	Yes	Yes	Yes

Tab. 2. A comparison of usability evaluation methods for e-learning systems

(Table 2 continued)

Criteria	e-Learning usability evaluation methods					
	SUE methodology	**eLSE methodology**	**MiLE+ method**	**Ssemugabi & de Villers´sframework**	**Zaharias´s Questionnaire**	**Nokelainen´s PMQL**
Evaluation target	Platformandmodule	Platformandmodule	Platform	Platform	Platform	Platformandmodule
Evaluation of stake holders' roles	Experts – no different roles Users – students only	N/A	Experts – no different roles Users – students and instructors	Experts – perform different users roles Users – students only	Users – employees only	Users – students only
Empirical evidence of the method	Yes, 1 pilot study	Yes, 1 comparison study	Yes, several	Yes, pilotstudy	Yes, 2 pilot studies and questionnaire validation	Yes, 2 pilot studies and questionnaire validation
Empirical comparison to other methods	No	Yes, AT inspection compared to adapted HE and thinking-aloud	No	Yes, adapted HE against questionnaire	No	No
Future developments of the method	Indicated	Not indicated	Indicated	Indicated	Indicated	Not indicated
Applicably in phase	N/A	N/A	N/A	N/A	N/A	N/A
Required time	N/A	N/A	High (21 persons/day in 6 months)	N/A	Low*	Low*
Required budget	N/A	N/A	N/A	N/A	N/A	N/A
Needed users	10 (pilot study)	N/A	12 students + 9 instructors	High (more than 60)	High (more than 100)	High (more than 50)

(Table 2 continued)

Criteria	e-Learning usability evaluation methods					
	SUE methodology	**eLSE methodology**	**MiLE+ method**	**Ssemugabi & de Villers´sframework**	**Zaharias´s Questionaire**	**Nokelainen´s PMQL**
Required evaluators	N/A	N/A	2	4	Not aplicable	Not applicable
Required expertise	N/A	Evaluatorswith limited experience	Noviceevaluators	Medium to high (2 double experts)	Not aplicable	Not applicable
Required equipment	N/A	N/A	Medium*	N/A	Low*	Low*
Method in trusiveness	Partially (during user testing)	Partially (during user testing)	Partially (during user testing)	No	No	No
Style of evaluation	Laboratory study*	Laboratory study*	Laboratory study*	Laboratory/ field study	Field study	Laboratory/ field study
Level of subjectivity or objectivity of the technique	Both subjective and objective*	Boths ubjective and objective*	Both subjective and objective*	Both subjective and objective	Subjective	Subjective
Type of measures provided	Qualitative	Qualitative	Qualitative with quantitative data	Qualitative and quantitative	Quantitative	Quantitative
Information provided	High level	High level	High (extensive reports from evaluators)	Medium to high level*	Low to medium level*	Low to medium level*
Immediacy of the response	Partially (when thinking-aloud and interview is applied)	N/A	Partially (when thinking-aloud is applied)	No	No	No

5. Conclusions and future work

In this paper several methods for usability evaluation of e-learning systems that have emerged lately are compared. The comparison is done by identifying general and specific criteria which facilitate the selection of the appropriate method to determine usability problems. The selection of usability evaluation methods is not an easy task and is influenced by time, cost, efficiency, effectiveness, and ease of application (Ssemugabi & de Villers, 2009), as well as the scope of method application in the e-learning context. From the comparison of current e-learning UEMs by criteria that are crucial when choosing the appropriate research method, it is evident that many of them lack basic instructions about practical application and resources needed to perform the method. The lack of such information could prevent wider adoption of the method in practice, both in academic and professional communities.

Methods for usability evaluation of e-learning applications, such as SUE, MiLE+ or eLSE, which integrate several traditional usability evaluation methods, to some extent evaluate specific aspects of an e-learning platform or educational content. While their focus is on scenario-based usability inspection, user testing is not obligatory but may be driven for critical scenarios/tasks identified by evaluators. Evaluation of user's satisfaction or motivation is not performed.

On the other hand, approaches that are based on user testing integrate pedagogical usability into psychometrically validated questionnaires. However, these one-dimensional approaches are based on subjective assessment of users and lack identification of objective usability problems that are revealed with other methods.

So far, none of the examined methods has enabled comprehensive usability evaluation of e-learning platforms and educational modules considering a wide range of specific e-learning attributes. Thus, further research is needed to adapt the current methods to more integrative approaches. Without adjusting the current design guidelines and usability evaluation methods to the e-learning perspective, there is a danger that in usability studies examining e-learning platforms and e-courses identification of important usability issues, particularly pedagogical ones, will be omitted.

In order to address limitations of the existing UEMs for e-learning, authors of this paper are currently focusing their research efforts on developing a broader conceptual framework with parameters and heuristics for technical and pedagogical usability evaluation of e-learning systems and e-courses. That framework will provide a basis for the adaptation of several inspection usability methods and user testing methods to the e-learning context, forming an integrative usability evaluation method for e-learning. The new method will be validated on several e-courses provided on learning management systems.

6. References

Albion, P. (1999) Heuristic evaluation of educational multimedia: From theory to practice. Paper presented at *the 16th Annual conference of the Australasian Society for Computers in Learning in Tertiary Education*.

Ardito, C., De Marsico, M., Lanzilotti, R., Levialdi, S., Roselli, T., Rossano, V., and Tersigni, M. (2004) Usability of E-learning tools. In *Proceedings of the Working Conference on Advanced Visual interfaces* (Gallipoli, Italy, May 25 - 28, 2004). AVI '04. ACM, New York, NY, 80-84

Ardito, C., Costabile, M., Marsico, M., Lanzilotti, R., Levialdi, S., Roselli, T., &Rossano, V. (2006) An approach to usability evaluation of e-learning applications. *Universal Access in the Information Society*, 4 (3), pp. 270-283

Bolchini, D., Garzotto, F. (2008) Quality and Potential for Adoption of Web Usability Evaluation Methods: An Empirical Study on MILE+, *Journal of Web Engineering*, 7(4), pp. 299-317, Rinton Press

Costabile, M.F., De Marsico, M., Lanzilotti, R., Plantamura, V.L., Roselli, T. (2005) On the Usability Evaluation of E-Learning Applications. *Proceedings of the 38th Annual Hawaii International Conference on System Sciences,* HICSS '05.

Dix, A., Finlay, J., Abowd, G., Beale, R. (2004) *Human-Computer Interaction.* (Third Edition). Harlow Assex: Pearson Education Limited

Granić, A. (2008) Experience with usability evaluation of e-learning systems. In: Holzinger, A. (Ed.). *Universal Access in the Information Society*. Vol. 7, No. 3, pp. 209-221

Hollingsed, T., Novick, G.D. (2007) Usability Inspection Methods after 15 Years of Research and Practice. *Proceedings of the 25th Annual International Conference on Design of Communication, SIGDOC 2007*, El Paso, Texas, USA, October 22-24, pp. 249-255

Holzinger, A. (2005) Usability engineering methods for software developers. *Communications of the ACM*, Vol. 48, No. 1 (Jan. 2005), pp. 71-74

Hornbæk, K. (2006) Currentpractice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies*, Volume 64, Issue 2, pp. 79-102

Inversini, A., Botturi, L., Triacca, L. (2006) Evaluating LMS Usability for Enhanced eLearning Experience. In E. Pearson & P. Bohman (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2006*. Chesapeake, VA: AACE. pp. 595-601

ISO 9241-11 (1998) *Ergonomic requirements for office work with visual display terminals (VDTs)— Part 11: Guidance on usability*

Kothari, C.R. (1990) Research Methodology: Methods and Techniques (Second Revised Edition). New Age International Publisher

Kukulska-Hulme, A., Shield, L. (2004) Usability and Pedagogical Design: Are Language Learning Websites Special?. *ED-MEDIA'04 - World Conference on Educational Multimedia, Hypermedia and Telecommunications*, Lugano,

Switzerland, 22-26 June 2004. Published in the Journal of Educational Multimedia and Hypermedia, 15 (3), pp. 349-369

Lanzilotti, R., Ardito, C., Costabile, M.F., De Angeli, A. **(**2006) eLSE Methodology: a Systematic Approach to the e-Learning Systems Evaluation**.** *Journal of Educational Technology & Society***,** Vol. 9, Issue 4.

Levy, Y. (2007) Comparing dropouts and persistence in e-learning courses. *Computers & Education*, Volume 48, Issue 2, February 2007, pp. 185-204

Nielsen, J. (1994) Heuristic evaluation. In: Nielsen, J., Mack, R.L. (eds) *Usability Inspection Methods*. Wiley & Sons, New York

Nokelainen, P. (2006) An empirical assessment of pedagogical usability criteria for digital learning material with elementary school students. *Educational Technology & Society*, 9 (2), pp. 178-197

Preece, J., Rogers, Y. & Sharp, H. (2002) *Interaction Design: Beyond Human-Computer Interaction.* New York: Addison-Wesley.

Reeves, T.C., Benson, L., Elliott, D., Grant, M., Holschuh, D., Kim, B., Kim, H., Lauber, E., Loh, S. (2002) Usability and Instructional Design Heuristics for E-Learning Evaluation. In P. Barker & S. Rebelsky (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2002*, pp. 1615-1621. Chesapeake, VA: AACE.

Ssemugabi, S., de Villiers, R. (2009) A comparative study of two usability evaluation methods using a web-based e-learning application. In *Proceedings of the 2007 Annual Research Conference of the South African institute of Computer Scientists and information Technologists on IT Research in Developing Countries*, SAICSIT '07, Vol. 226. ACM, New York, 132-142

Silius, K., Tervakari, A-M.,Pohjolainen, S. (2003) A Multidisciplinary Tool for the Evaluation of Usability, Pedagogical Usability, Accessibility and Informational Quality of Web-based Courses. *The Eleventh International PEG Conference: Powerful ICT for Teaching and Learning*, 28 June - 1 July 2003, St. Petersburg, Russia.

Squires, D., Preece, J. (1999) Predicting quality in educational software: Evaluating for learning, usability and the synergy between them. *Interacting with Computers*, 11(5), pp. 67- 483

Triacca, L., Bolchini, D., Botturi, L., Inversini, A. (2004) MiLE: Systematic Usability Evaluation for E-learning Web Applications. In L. Cantoni& C. McLoughlin (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2004* Chesapeake, VA: AACE. (pp. 4398-4405). Conference ED Media 04, Lugano, Switzerland

Zaharias, P. (2006) A usability evaluation method for e-learning: Focus on motivation to learn. In *Extended Abstracts of Conference on Human Factors in Computing Systems - CHI2006*

Zaharias, P. (2009) Usability in the Context of e-Learning: A Framework Augmenting 'Traditional' Usability Constructs with Instructional Design and Motivation to Learn. *International Journal of Technology and Human Interaction (IJTHI)*, Vol. 5, No. 4, pp. 37-59