MiLE: Systematic Usability Evaluation for E-learning Web Applications

Luca Triacca, Davide Bolchini, Luca Botturi, Alessandro Inversini
University of Lugano, Switzerland
{triaccal, bolchind, botturil, inversia@lu.unisi.ch}

Abstract: This paper presents a proven and reusable methodology (MiLE) for performing a cost-effective usability evaluation of an e-learning web application. MiLE is a scenario-driven inspection technique which is based on the concepts of user profile, user goal, scenario, and usability attribute. Mitigating the drawbacks and merging the respective benefits of state-of-the-art methods for usability evaluation, MiLE is intended to be a helpful tool for project managers, instructional designers, and evaluators to carry out a learner-centered validation which can anticipate and analytically justify the usability breakdowns, thus providing organized indications for a focused redesign. Examples of the results that can be obtained using MiLE are showed through a real case study evaluation of a large e-learning corporate platform.

1. Introduction and Motivation

The main goal of usability evaluation is to detect the most part of the problems, obstacles and breakdowns for the user when interacting with a web application, being the usability "the effectiveness, efficiency and satisfaction with which specified users can achieve specified goals in particular environments" (ISO 9241-11). For e-learning environments and applications, usability is a necessary (although not sufficient) condition for effective online learning. Easily locating and accessing the needed content, orienting oneself in the maze of different paths and nested pages of a structured website, avoiding being overloaded by the information clustered in a page, and being able to use effectively the navigation architecture are just some examples of important conditions for a learner to accomplish his/her learning tasks.

In other words, usability is the property of a mediated learning environment of supporting the users as transparently as possible in the accomplishment of their learning goals.

It has been acknowledged that hypermedia contents are more and more a great sum of information but they are being poorly structured (Acosta et al. 2003). If we assume that the organization and the usability of the information is partly responsible for a better learning (Najjar, 1996), then a usable e-learning website is not just a resource with a nice "look&feel", but a web application which communicates contents and structures the interaction in such a way that facilitates the learning experience. However, it is clear that usability evaluation is just one specific aspect the quality assessment of a e-learning environment. A highly usable online course does not guarantee at all a high quality in the learning outcome. Nevertheless, shaping usability represents an important condition for success of e-learning projects.

How can usability of an e-learning web application be effectively evaluated and measured? How can evaluators, instructors, reviewers, tutors and course managers be equipped with usable and ready-to-use tools for evaluating the usability of the web application prototyped or deployed?

This paper presents an adapted version of a proven methodology (MiLE - is the acronym for Milano-Lugano Evaluation Method. It is the result of a long-standing cooperation between Politecnico di Milano and University of Lugano.) for usability evaluation which has been already extensively and successfully used in a variety of web application domains (e.g. educational institutions, cultural-heritage, public education, and e-government) (Matera et al., 2002) (Bolchini et. al., 2003) (Triacca et.al., 2003), and that has been used and tailored for e-learning applications. MiLE is valid for the usability evaluation of e-learning interactive applications and, within the specific case presented, it showed how MiLE has been adopted and adapted for the evaluation of an e-learning web-based system. The goal of the contribution is to provide course developers and instructional designers with a structured "kit" of guidelines and practical suggestions for a cost-effective usability evaluation of their online application. With respect to the commonly used guidelines and checklists available in the literature, the expected benefits coming from the adoption of MiLE are a better anticipation of the usability breakdowns that may jeopardize a satisfactory learning experience, as well as a more analytic provision of feedback for redesign.

The paper is organized as follows: section 2 gives an overview of the main approaches and techniques used for usability evaluation of web-based applications. Section 3 explains the essentials of MiLE through concrete examples

from the e-learning domain. Some of the results of the application of MiLE to a corporate learning platform of an internationally-renowned car manufacturing company are presented in section 4. Finally, concluding remarks and reflections on the experience of the case study are followed by suggestions for future research (Section 5).

2. Related work

Over the last decade, HCI practitioners and researchers have defined and assessed a variety of approaches and methods for evaluating the usability of information systems. These usability approaches has been increasingly improved and refined from their application to traditional information systems to their wide spread use for web-based applications. The most commonly adopted approaches to web usability are user-based methods (or user-testing methods) and usability inspection methods (or expert reviews) (Matera et al., 2002).

User-based methods mainly consist of user testing, in which usability properties are assessed by observing how the system is actually used by some representatives of real users (Whiteside Bennet & Holtzblatt, 1988) (Dix A. et. al., 1998). User-testing evaluation provides the trustiest evaluation, because it assesses usability through samples of real users. However, it has a number of drawbacks, such as the difficulty to properly select correct user samples and to adequately train them to manage advanced functions of a website (Matera et al., 2002). Furthermore, it is difficult, in a limited amount of time, to reproduce the actual usage situation. This condition is called the "Hawthorne effect" (Roethlisberger & Dickson, 1939): observed groups can be affected by observation alone. Failures in creating real-life situations may lead to "artificial" conclusions rather then realistic results (Lim, Benbasat & Todd, 1996). User testing is considerable in terms of time, effort and cost. However, it is effective to evaluate quickly the look and feel of the interface, as it is possible to verify at "real-time" the reactions of the users.

Usability Inspections methods is the generic name for a set of methods based on having expert evaluators analytically examine usability-related aspects of a user interface (Nielsen & Mack, 1994). With respect to user-testing evaluation, usability inspection is more subjective, having heavy dependence upon the inspector's skills (Matera et al., 2002). The main advantage of the inspection methods is the relationships between costs and benefits. As a matter of fact, performing an inspection can "save users" (Nielsen & Mack, 1994), (Jeffries et al. 1991) and does not require any special equipment. The inspector alone can detect a wide range of usability problems and possible faults of a complex system in a limited amount of time (Matera et al., 2002). For these reasons, inspection methods have achieved widespread use in the last years, especially in industrial environments (Madsen, 1999).

Current usability inspection methods have two main drawbacks. Firstly, they focus on "surface-oriented" features of the graphical interface (mainly at page level) (Green & Benyon, 1996). Only few of them address the usability of the application structure, e.g., the overall information architecture, organization of content or navigation patterns. Secondly, the reliability of the results is often entirely dependent on the individual know-how, expertise and skills of the inspectors.

Both user-based methods and inspections methods are alternatively based on two techniques: heuristic-driven evaluation and task-driven (or scenario-based) evaluation.

Essentially, heuristic-driven evaluation provides checklists and usability principles (Nielsen, 1999). The quality of the website is assessed against these principles (e.g. consistency, reliability, status visibility, etc.). In user testing, heuristics are used to ask users to comment in a structured way their experience with the website (e.g. heuristic questionnaires or structured interviews). During inspection, heuristics guide the expert to explore the site and check compliance with usability principles.

Although traditional heuristics have been already extended and specified for coping with some distinctive features of the e-learning applications (Reeves, 2002), the instructional design and e-learning communities still lack a widespread culture of usability. This is partly due to the fact that a lot of educational web application development is often constrained by existing development platforms as WebCT or BlackBoard – thus leaving a great part of design to the platform structure.

Task-driven evaluation assumes that usability is assessed by trying to complete actions with the website. Tasks are provided which describe potential goals or sequences of actions that users might want accomplish with the application. In user testing, tasks are defined and given to the users. Tasks are also employed in walkthrough and other structured inspection techniques (Rosson & Carroll, 2002) (Brinck, Gergle & Wood, 2002).

Heuristics and task-driven techniques are usually adopted in alternative and separately, thus loosing the opportunity to obtain a more comprehensive evaluation.

Moreover, one of the main disadvantages shared both by heuristics and task-based techniques is that they are not reuse-oriented, i.e. they have not been defined to be effectively reused. Most of usability techniques are proprietary methods or expert-dependent techniques; in other words, they can be difficult for less-experienced evaluators who

do not have the necessary conceptual tools to gain appreciable results. The problem of reuse is strongly connected to the difficulty of teaching and communicating the essence of a method in a way that other people can successfully apply it. Projects teams are acknowledging the importance of usability evaluation but are still reluctant to make considerable investment in consultancy for an "ad-hoc" evaluation, especially if the development phase is ending and the remaining budget is very limited. Effective reuse of usability knowledge and practices would enhance the adoption of usability techniques by designers and project teams.

3. MiLE: blending inspection-driven scenarios with heuristics for e-learning usability

MiLE is an experience-based evaluation framework for web applications that strikes a healthy balance between heuristic evaluation and task-driven techniques. MiLE offers reusable tools and procedures to carry out both expert reviews (also called inspections) and user testing within budget and time constraints. Basically, MiLE uses scenario analysis as the driver for usability evaluation. Scenarios are then combined with selected usability heuristics in order to provide analytic feedback on the different aspects of the applications design.

Assuming a scenario-based perspective, without a clear understanding of the need of the learners and their goals, it is really hard to perform a usability evaluation that may provide useful and in-depth results. To this end, the knowledge of the domain and of the context of use of the application will surely facilitate the evaluators in this task. We can assume that in a typical e-learning application there are at least two main categories of stakeholders who interact with the website: the students and the instructors (everyone with particular needs and goals). Each category uses a specific part of the website. On one hand, the students use the front-end for gathering and learning the course materials and interacting with the tutors and colleagues. On the other hand, the instructors use mainly the back-end of the application for uploading course documents and managing the interactions, as well for organizing, updating and tuning the course instructional architecture.

In this context, the first thing to do in order to prepare a usability evaluation is to identify the user profiles and their goals. In fact, at a deeper analysis, the simple distinction of types of users in students and instructors may actually give raise to a more refined definition of the user profiles considered relevant for the evaluation. Profiles identify specific roles that may be played a generic user type (Table 1).

User Type	
	User profiles
Learner	Responsible
	Accepter
	Seller
Instructor	Tutor (1..n)
	Administrator
	Responsible

Table 1. Structuring User Profiles.

Users have generic goals they would like or they are supposed to accomplish through the web application, being clear that these goals are strictly related to the set of the overall learning goals of the course. Therefore, it is possible to associate one or more high-level goals to each user profile, thus creating the essential constituents of a user scenario. Since we are here considering very high-level (or macro) goals, we will call these artifacts macro-scenarios (Table 2).

Macroscenario A	
User profile	**Seller**
Macrogoal	**Plan the learning experience**

Table 2. An example of macroscenario.

A typical line of inquiry for eliciting macro-goal may be summarized in the question: "What may a user profile want to use the e-learning website for?"

Macro-scenarios capture a general target of achievement, which may be accomplished through several strategies or subgoals. Evaluators may thus identify more detailed scenarios that should take place in order to accomplish a given macro-scenario. This refinement process is usually led by questions like: "How may the learner accomplish the

macro-goal? What should s/he be able to do to get to the macro-goal?. In this way, evaluators define new lower-level scenarios, which try to anticipate in a structured and organized way the expected user experience (Table 3).

Macroscenario A	Plan the learning experience
Scenario 1	Plan the study
Scenario 2	Know course conditions
Scenario 3	Know the learning level achieved

Table 3. Refining macro-scenarios into scenario.

Now it is possible to define a series of *tasks* which describe the activities the user should perform on the website for each goal identified. Note that this refinement process from high-level goals to detailed tasks is not intended to be complete and exhaustive. Evaluators should identify the critical goals and tasks which they hold as important – according to their instructional experience – for the learners to accomplish each macro-scenarios.

Macroscenario A	Plan the learning experience	
Scenario 1	Goal	Tasks
	Plan the study	− Know the time required to frequent a course − Find the ideal period to frequent a classroom session − Know the time needed to download a document
Scenario 2	Know course conditions	− See course goals − See the course structure − See how to communicate with tutors and peers
Scenario 3	Know the learning level achieved	− Make a test in order to verify the level of learning achieved − Verify in which topic there are gaps

Table 4. Refining scenarios into user tasks.

Finally, the result of the scenario definition is a structured set of tasks (and relative goals) associated to each user profile (Table 4). The next step for the evaluator is trying to perform each single task on the application and evaluate *whether* and *how* they are feasible.

While performing each task by traversing and browsing the pages and links, evaluators do not only assess how effectively and efficiently they can complete the task, but they are also supported by specific heuristics that guide the inspection to focus on the different application aspects that are relevant for the user experience, such as the content offered, the effectiveness of the navigation, the clarity of the link labels, and the orientation clues within the information architecture.

Application Aspect	Usability Attribute	Description
Content	Authority	The author is competent in relation to the subject
	Currency	The time scope of the validity of the information is clearly stated. The info is updated
	Completeness	The information required is sufficient to complete the task
	Richness	The information required is rich (e.g. enough examples)
	Clarity	The information is easy to understand
	Multilinguisticity	The information is given in more than one language
Structure of Content	Structure effectiveness	The organization of the content pieces is suitable to the task
	Consistency	Similar pieces of information are organized with in similar fashion
Navigation	Accessibility	The information is intuitively located and easily accessible
	Orientation	It is easy to understand where I am within the information architecture
	Link promises	Links actually lead to the content they promise to lead
Interface & Presentation	Multimediality	Different media are used to convey the information necessary to complete the task
	Predictability	Link labels spread enough "scent of information", i.e. it is clear where they lead to
	Visual Communication	The quality of the visual design as how it conveys effectively content and interaction capabilities

Table 5. Examples of usability attributes.

These different dimensions of the web application have been organized in a library of usability attributes, which evaluators may pick at their choice taking into account the purpose and constraints of the task at issue. Table 5 shows a sample list of usability attributes usually employed by MiLE inspectors.

A subset of these attributes may be considered by the inspector in order to give an analytic evaluation on a specific application aspect in the context of executing a task.

In other words, as the inspector is performing that task "*See the course structure*", s/he will be attentive as to how s/he can get easy orientation during the path, as to how link labels offered are predictable, as to how the information about the course structure is easy to locate and to reach, as to how the structure of the course is presented in a clear and satisfactory way, and so on.

Being MiLE a method that can be used flexibly at different levels of granularity – according to the resources available to the evaluators – usability attributes may be considered not necessarily only when performing detailed tasks. Inspectors may evaluate even goals and macrogoals using a subset of the usability attributes (Figure 1). In this case, it is important to note that the time and resources saved has to be balanced with a coarser-grain analysis.
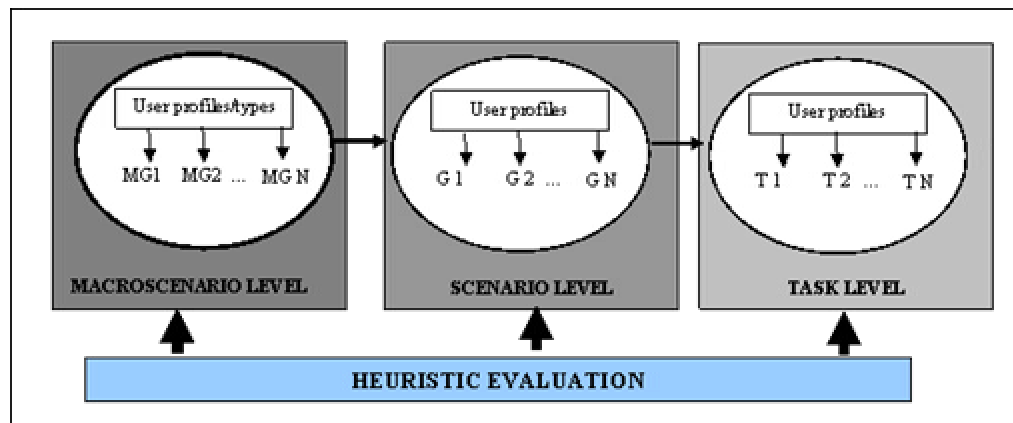

Figure 1. Using heuristics to analytically evaluate each scenario at different levels of detail
(MG: Macrogoal, G:.Goal, T: Task).

Although originally developed for hypermedia web application (not specifically for e-learning), MiLE is now attracting increasing interests for its application to e-learning environments. It is clear that the suite of concepts provided by MiLE (user types and profiles, user goals, macro-scenarios, scenarios, tasks, and usability attributes) and the methodology to use them represents a powerful evaluation tool also for educational web applications. To start validating this claim, we had the opportunity to be committed in a real usability evaluation for the e-learning corporate platform of the Italian branch of an internationally renowned car manufacturing company. In this context, we used MiLE throughout the usability evaluation, thus gathering analytic results that were reported to the platform managers as requirements for specific interventions of redesign.

4. Making use of MiLE on an e-learning corporate platform

The e-learning community center of the car manufacturing company at issue aims at providing a support for follow up to classroom-based activities (e.g. material distribution), as well as at providing a quick and always up-to-date information channel for accepter and sellers of the company distribution system, through the delivery of short self-contained "learning pills". Finally, one of the goals of the application was to offer an engaging way of training for involving accepter and sellers with the corporate training system.

After the creation of the macro-scenarios thorugh interviews and focus groups with the stakeholders (course responsible and other consultants), we carried out usability evaluation mainly at the task level, using the usability attributes that were relevant for each task for specifying the different problems encountered. The result of the usability evaluation analysis is a report that presents the complete scenarios analysis (in total 4 macroscenarios and 8 scenarios have been analyzed for the user type "students", 3 macroscenarios and 7 scenarios for the type "instructor"). Here we present some examples of the results obtained. The overal usability evaluation took 21 men/day over a period of 6 months.

4.1 Examples of front-end usability problems

When performing the task *See the course structure* within the goal *Know course conditions*, we found out a clear outline of the course organization is missing. An outline of the course structure would help the learners effectively plan the study in detail, by understanding the needed time to complete a course, and thus allocating the needed resources for it. A suggestion for redesign would be to specify and highlight the organizations of the learning units (e.g. modules and sub-modules) within a course, preferable in chronological order.

Within the goal *Plan the study*, when trying to accomplish the task *Find the ideal period to frequent a classroom session* (each person can decide when to take a course among a set of dates), a problem (which has been encountered also in other tasks) that could affect the student experience is the poor self-evidence and visibility of textual links. For this reason, the suggestion is to create a consistent style to be used for all textual links, so that the user recognizes immediately when and where a textual link is present (Figure 2).

Consistency: 2 different visual styles for textual links.

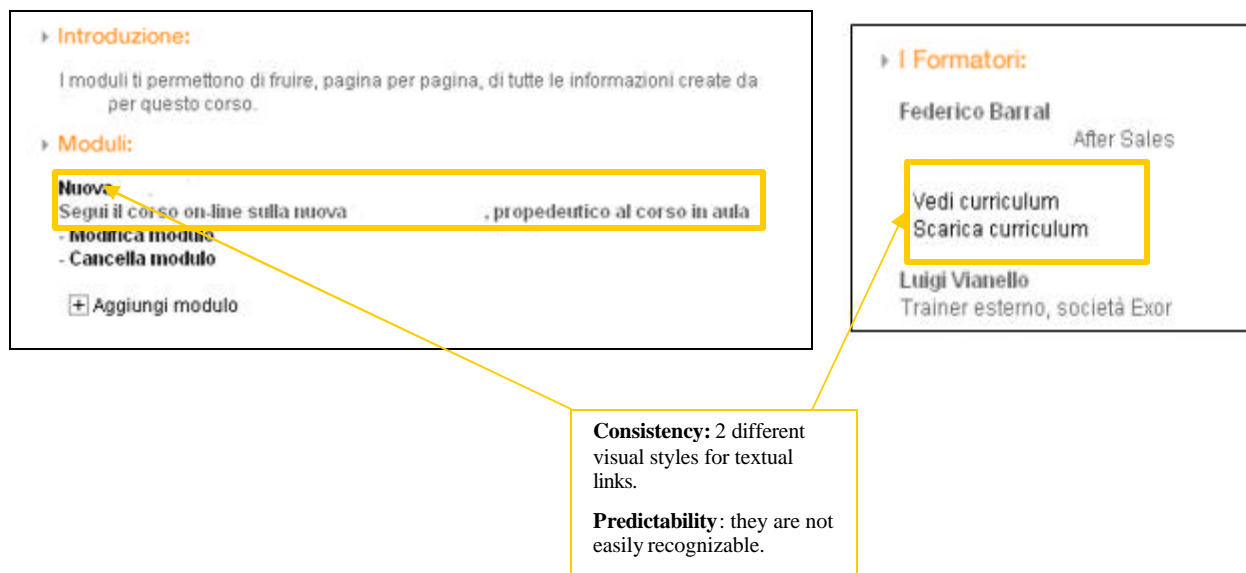**Predictability**: they are not easily recognizable.

Figure 2. Consistency and predictability problems of link visibility.

Another usability problem is related with the main menu (Figure 3). In the context of the task *See course structure*, it was noticed by the evaluators that the orientation is poor: it is not clear which sub section the user is currently browsing. It would be useful to highlight (by means of colours or graphic styles) the current position of the user within the suite of courses offered.
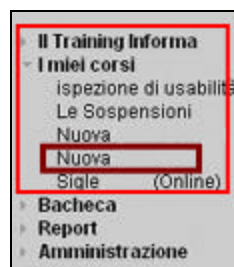

Figure 3. Orientation breakdown during the task "See Course Structure"

The third main usability problem is that in some cases the user cannot immediately understand the consequences of his/her actions. For example, when the user tries to download a file there is no explanation of what type of file she/he is downloading, what its size is. Within the scenario *Study a course* and, more specifically, when doing the task *Find course materials seen in classroom session*, it can be noticed an interesting differentiation between classrooms materials and generic materials. What could that mean? Here we have a predictability problem.

Moreover, this distinction is not always used for every class course (consistency problem). To improve the usability of this aspect of the application we suggested to use the label "online material" for materials provided only by means of the application (e.g. videos) and to add another category of materials labelled "classroom material" (in this case the user downloads the material which has been handed out in classroom) . The goal of both actions is to orientate the user in the different kinds of materials she/he can find in the application.

4.2 Examples of back-office usability problems

During the instructor macroscenario *Manage a course*, the main usability problems identified for the back-office are related to the orientation when an instructor browses the different sections.

In particular, there are two different sections called "administration". For this reason it is very difficult to understand immediately the difference between them. The administrator (especially if not yet familiar with the application) may feel confused seeing two sections with the same name and different functionalities.
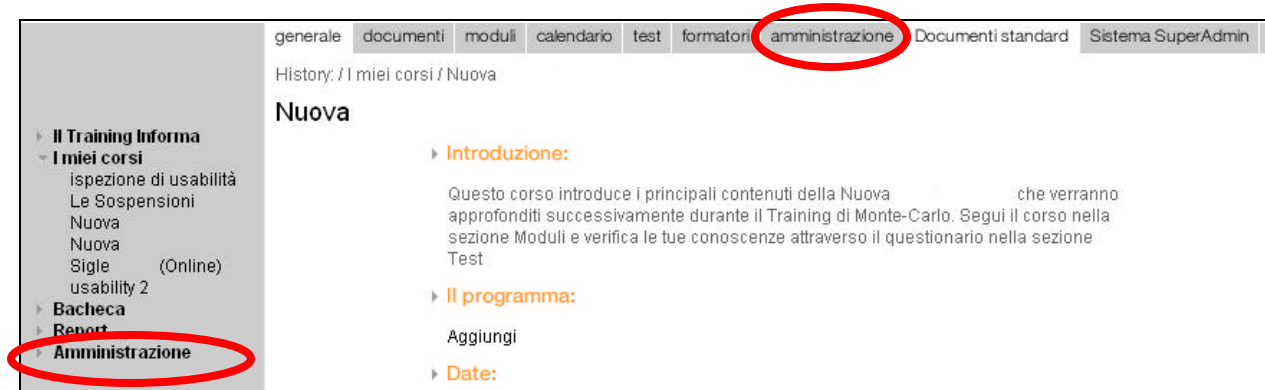

Figure 4. Consistency in link labeling is infringed. This may obstacle orientation.

The contents of these sections are very different. We foresee that it would be real hard for the administrator to recognize the difference between the two different administration sections. It should be better to have a common semantic for the entire the back-end side (Figure 4).

The suggestion is to label differently the sections (for example "Amministrazione generale" – general administration - and "Amministrazione corso" – course management -. In this case, it is very important to use a common semantic for navigation and content. The administrator should know before trying to click a text string if it is really a link or not. Also, the administrator should know the consequences of his actions.

When performing the task *Build a test*, the instructor is faced with the possibility of adding a question to the test. However, the command that the instructor is supposed to click to confirm the will of adding a question is "Avanti" ("forward") (Figure 5).
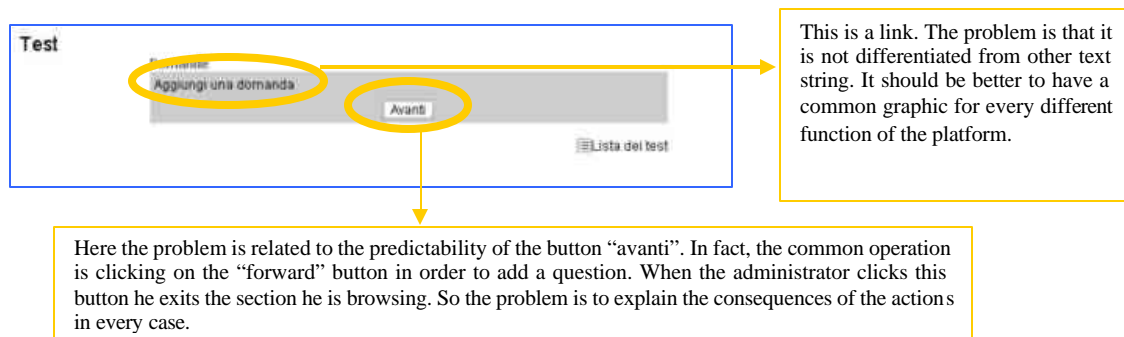

Figure 5. Predictability problems in the task "Build a test".

These and other problems were identified using MiLE. Scenario analysis helped discovering critical areas of the application that inspector consider potential usability problems. Using specific heuristics while performing scenarios facilitated the definition of the problems and enhanced the communication of the issues to the designers.

5. Conclusions and Future Work

We presented a systematic and practical approach (MiLE) for evaluating the usability of an e-learning application, taking into account the user requirements, their goals and scenarios of use. MiLE is based on inspection and combines scenario-based evaluation (when inspectors verify the feasibility of given tasks on the application) with heuristic-driven evaluation (which verify the compliance of the website with a set of usability principles). E-learning applications may greatly benefit from such a structured technique because specific feedback to the designers may be provided on the basis of the results of the inspection. User testing may also be organized and performed in order to confirm, validate and mitigate the opinions of the inspectors, and to find new usability breakdowns that might have been overlooked.

Future work will consolidate the results of the usability inspections in order to foster technology transfer of MiLE to instructional designers and online course managers active in organizations and higher-education. One issue at stake here is how usability can be integrated as a concern in the instructional design process, and not be relegated to an ex-post analysis. Moreover, we are going to explore relevant e-learning domains for which we are able to gather substantial evaluation experience using MiLE, in such a way that a reusable repository of recurrent user profiles, macro-scenarios and combination of usability attributes relevant for a given e-learning domain (e.g. corporate training or higher education, etc.) will be incrementally defined. Another interesting element is providing a general usability evaluation of the most widely used platforms, suggesting then key actions that designers can take in order to enhance their courses when supported by those platforms.

References

Acosta M.P., Monguet J.M., Rodriguez R. (2003). Educational Hypermedia Applications: Design Based on Content Models, ED-MEDIA 03, Honolulu, Hawaii, USA.

Bolchini D., Triacca L., Speroni M. (2003). MiLE: a Reuse-oriented Usability Evaluation Method for the Web, HCI International Conference, Crete, Greece.

Brinck, T., Gergle, D., Wood, S.D. (2002) Usability for the web, Morgan Kaufmann.

Carroll, J. (2002), Making Use – Scenario-based design of Human-Computer Interactions, MIT Press.

Cato, J. (2001). User-Centred Web Design, Addison Wesley.

Di Blas et al. (2002). Evaluating The Features of Museum Websites, Museum & Web Conference, Boston.

Green T.R.G., and D.R. Benyon. (1996). The skull beneath the skin; Entity-relationship modeling of information artifacts, int. J. Hum.-Comput. Stud., vol.44, no.6, pp.801-828.

Jeffries R., Miller J., Wharton C., and Uyeda K.M. (1991). User interface evaluation in the real world: A comparison of four techniques, in Proc. ACM CHI, pp.119-124.

Lim K.H., Benbasat I., Todd P.A. (1996). An experimental investigation of the interactive effects of interface style, instructions, and task familiarity on user performance, ACM Trans.Comput.-Hum. Interact., vol.3, no1, pp.1-37.

K.H. Madsen. (1999). Special issue on the diversity of usability practices, Commu nication ACM, vol.42, no.5.

Matera, M. et al. (2002). SUE Inspection: An Effective Method for Systematic Usability Evaluation of Hypermedia, IEEE Transaction, Vol.32, No. 1.

Najjar, L.J., Multimedia information and learning. Journal of Educational Multimedia and Hypermedia, 5, 129-150.

Nielsen, J., Mack R. (1994).Usability Inspection Methods, Wiley.

Nielsen, J. (1999). Designing Web Usability, New Riders.

Reeves, T.C. (2002). Usability and Instructional Design Heuristics for E-Learning Evaluation, EDMEDIA 2002, Denver, USA.

Reeves, T.C., Carte B.J. (2001), Usability Testing and Return on Investment Studies: Key Evaluation Strategies for Web-based Training, in B. Kahn (Ed.), Web-Based Training. Englewood Cliffs, NJ: Educational Technology Publications.

Rosson, M.B., Carroll, J. (2002). Usability Engineering, Morgan Kaufmann.

Roethlisberger, F.J. and Dickson,W.J. (1939). Management and the Worker, Boston, Mass.: Harvard University Press.

Triacca L., Bolchini D., Di Blas N., Paolini P. (2003). Wish you were Usable! How to improve the Quality of a Museum Website. International Conference on Electronic Imaging and the Visual Arts (EVA03) Florence, Italy.

Whiteside J., Bennet J., and Holtzblatt K. (1988), Usability engineering: Our experience and evolution, in Handbook of Human-Computer Interaction, M.Helander, Ed. Amsterdam, The Netherlands, North-Holland.