
Towards the evaluation of usability in educative websites

Ma. Elena Alva*, Ana Belén Martínez,
Ma. del Carmen Suárez, José Emilio Labra,
Juan Manuel Cueva and Hernán Sagástegui

Department of Computing of the University of Oviedo,
C./Calvo Sotelo, s/n. 33007 Oviedo, España

E-mail: alvamaría@uniovi.es

E-mail: belenmp@uniovi.es

E-mail: macamen@uniovi.es

E-mail: labra@uniovi.es

E-mail: cueva@uniovi.es

E-mail: hsagastegui@uniovi.es

*Corresponding author

Abstract: This article presents a methodological approach of evaluation of the usability in educational environments ME-UsitE, an analysis of the necessity to implement a tool of evaluation and measurement of the usability to support the methodology, that contributes to the detection of problems and failures of usability from the perspective of the user of educative websites in a real work environment. Taking into account that the websites are designed and implemented for a domain and audience in particular, it is important to consider factors like these for the implementation of a tool that allows not only to carry out the evaluation process but also to support the analysis of the gathered data. Thus, the obtained results may allow us to establish policies for the improvement of the usability of the site, when providing a score that determines whether the level of usability offered covers the requirements for the proposed audience and the profile of the participant user in the evaluation.

Keywords: usability; educational; tools; evaluation parameters; measurement; methodology; user profile; level usability.

Reference to this paper should be made as follows: Alva, M.E., Martínez, A.B., Suárez, M.d.C., Labra, J.E., Cueva, J.M. and Sagástegui, H. (2010) 'Towards the evaluation of usability in educative websites', *Int. J. Technology Enhanced Learning*, Vol. 2, Nos. 1/2, pp.145–161.

Biographical notes: María Elena Alva Obeso graduated as an Industrial Engineer from Industrial Engineering School in 1986 at the University of Trujillo, Peru. She received her Master in Information System in 1998 at UPAO-Peru and I.T.S.M-Mexico and her PhD in 2005 at Oviedo University, Spain. She is a Professor of Programming Languages in the Engineering Faculty at the Antenor Orrego University in Peru from 1990–2005. She is a Professor at the Languages and Computers Systems Area at the University of Oviedo, Spain. Her research interests include web usability, web engineering and knowledge society.

Ana Belén Martínez Prieto graduated as a Computing Science Engineer in 1993 and received her PhD in Computing in 2001 from the University of Oviedo, Spain. She is an Associate Professor since 1995 at the Department of Computing of the University of Oviedo. Her research interests are new paradigms of databases, usability and accessibility.

María del Carmen Suárez Torrente graduated as Computing Science Engineer in 2000 from the University of Oviedo. She works as an Assistant Professor at the Department of Computer Science in the University of Oviedo since 1999. Currently she is working on her Doctoral thesis. Her research interests include web usability, web engineering and object-oriented programming methodology.

Jose Emilio Labra Gayo graduated as Computing Science Engineer in 1993 from the University of Oviedo. He received his PhD in Computer Science Engineering from the University of Oviedo in 2001. Since 2004, he is the Dean of the School of Computer Science Engineering at the University of Oviedo. His research interests are semantic web technologies, programming languages and web engineering, where he has published a number of papers in selected conferences and journals. He participates in several programmer committees like the *International Conference on Web Engineering*, the *International Workshop on Social Data on the Web* and the *Web Services and Service Oriented Applications Spanish Conference*.

Juan Manuel Cueva Lovelle graduated as Mining Engineer in 1983 from Oviedo Mining Engineers Technical School, University of Oviedo, Spain. He received his PhD in 1990 from the Madrid Polytechnic University, Spain. He is a Professor at the Languages and Computers Systems Area and since 2008 he is the Director of Computer Department in the University of Oviedo, Spain. His research interests include object-oriented technology, human-computer interface, object-oriented databases, web engineering, object-oriented languages design and object-oriented programming methodology.

Hernán Sagástegui Chigne graduated as Mathematician in 1990 from Sciences School, University of Trujillo, Peru. He received his Master in Computing Science in 1993 from the University of Cantabria, Spain. He was a Professor at the Mathematical School in the University of Trujillo in Peru in 1985–1992. Since 1991, He is a Professor of Programming Languages in the Engineering Faculty of the Antenor Orrego University, Peru. Currently he is working on his Doctoral thesis in the University of Oviedo. His research interests include web usability, knowledge society, e-Learning and collaborative learning.

1 Introduction

Though it is true that it have been achieved advances in the development of the applications centred on the user with the emergence of new technologies, also it is that in many cases they limit the access to great users diversity if we consider the necessary requirements of hardware and software to access them. On the other hand, although new developments are accessible, by no means they guarantee that they are more usable.

The educational websites are characterised for contributing at the entry of its users to the information society, allowing the universal access to the information and the knowledge without discrimination of race, colour, religion, origin or sex (Celik and Ipcioglu, 2007). Nevertheless, the complexity of the educational services offered in

internet, though they assure that the access to them should be sufficient, they do not assure the successful use of the content. Therefore, the study and evaluation of web usability are turning into important aspects to take into account in the development of web applications and, specially, in the educational domain.

The development of an educational web site is usually orientated to a dispersed user's community with diverse features. It is necessary not only to investigate the diversity of these users with regard to their skills, age, gender, culture, etc.; but also knowledge of these in order to reduce the gap between what the users know and what they need.

This study tries not only to establish a methodology capable of evaluating the usability of an educational website, but also to enrich web applications of other domains from this scheme of work. For example, the governmental web applications that offer services orientated to the user (a wide, complex and diverse community) and that therefore it needs to invest in research that allows to fix principles and guidelines to improve the level of usability of their sites and services, since this way can favour the integration of the different collectives, especially those with special needs, as well as the entrance of the companies in the society of the information.

1.1 Problems in the evaluation of the usability in educative websites

Many investigations have been directed to the activities of evaluation of the usability (methods, techniques, tools). Nevertheless, little it has been made on like integrating these activities in methodical and systematic coherent process, that allows the analysis of the usability results and that take into account the critical aspects in the evaluation as they are the user group (child, young, adult, etc.) or as it is the profile of these (inexperienced, intermediate, advanced).

At present time, different methods and techniques exist, that can be used during a usability evaluation: depending on the purpose of the measurement, the type of measurement to obtain, development cycle stage, etc. [investigation (Kirakowski et al., 1998), heuristics (Nielsen and Mack, 1994; Nielsen and Molich, 1990; Nakwichian and Sunetnanta, 2003), evaluation of experts (Schneiderman, 1998)]. Nevertheless, in their majority, these methods have been applied with, really, little participation of the real user in their real environment.

If we consider that the acceptance and the success of an educative website depend mainly on their capacity to cover the necessities of real users that they interact with this type of sites in private environment or at least real environment, it is important to consider the features of these users (abilities, interests, habits, age, among others.) like critical factors in the evaluation (Mineta, 2000; Nielsen, 2005).

Many are the authors whom they have investigated and proposed recommendations (guides, check lists, principles) for the development of usable applications (Nielsen, 1994; Tromp and Benford, 1997), but very little has been made on the usability evaluation of the educative websites (Carvalho, 2002). Nevertheless, it is enough to make an exploration in the web to verify that this still continues being a problem. As well pointed out by Murray and Constanzo (2000), the usability degree of a website depends on the intention and the objective audience.

Authors like Nielsen (2001), Human Computer Interaction Lab/University of Maryland (1998), Nakwichian and Sunetnanta (2003) have carried out studies of the usability of websites (in general), pointed out their applicability to web educative.

However these have been inferred considering that all the domains require the same aspects of usability or with the same priority.

1.2 Need of a methodology that systematises this process

Aspects as the indicated ones in the previous section limit the applicability of the existing tools, like support the evaluation of the usability in the interest domain. For that reason, it is necessary that the adoption of a methodology that includes the most suitable combination of the methods, techniques, etc., always focused to the profile of the user in the domain of the educative websites and that allows, of systematic way, to obtain reliable results.

This article shows the methodology of evaluation of usability of educational website: ME-USitE (Alva, 2005), which includes both the evaluation of the user and the evaluation of the expert as a more complete proposal to the evaluation in the approach of the usability. For the first one, there is a established general list of parameters of evaluation of the usability that includes both aspects of interface and aspects of content and basing on her, realises the selection only of those that fit to the needs of usability of a particular domain depending on the audience (child, young, adult and elderly) for which was designed and establishes the type of requirement, that is to say, if a requirement is essential (that is determinant in the usability of the site), desirable (that affects partially the usability) or optional (whose absence is not significant in the usability of the site). We will extend on ME-USitE in the following section.

1.3 Need of the tool of support to ME-USitE

Although it is true that the process of evaluation and measurement of the results of the proposal methodology can be carried out in a manual way also, it is true that we must think that the process of evaluation involves on one hand the application of questionnaires to users (for the obtaining of his/her profile as well as for the obtaining of the data of the evaluation) and the use of lists of verification on the part of the experts and on the other hand the necessity to process the data obtained thus to achieve concrete results with the consumption of resources and time that this implies. For that reason, it seems necessary to use automated tools to give support to the process of evaluation for the obtaining of data as well as for its treatment and processing.

At the moment, existing tools of support to the measurement of usability have been investigated and some have been oriented mainly to give support to the evaluation of commercial sites like WAMMI (Kirakowski et al., 1998), MUMS (Human Factors Research Group), QUIS (Su et al., 1987), PROKUS (Zülch and Stowasser, 2000), DRUM (Macleod and Rengger, 1993), etc.

This study has shown that these tools are oriented to measure the satisfaction and performance of the user or the product in web applications, but its development is focused on the domain of commercial applications. In addition, they consider that all the requirements of usability evaluation are same for any type of site, independently of the application and the audience, reason why its extension to other domains is possible according to Human Computer Interaction Lab/University of Maryland (1998) and Zülch and Stowasser (2000).

Nevertheless, each type of domain of application in the web has its own objectives of development and is designed for a specific audience (children, young, etc.) (Nielsen, 2005; Mineta, 2000), which can differ in the level of experience due indeed to the dispersed nature of mean of transmission of used knowledge (internet) that allows the access of different users with diverse motivations. Therefore, we think that it is advisable to design an appropriate tool to give support to the evaluation.

This article is structured the following way: in Section 2 a description of the methodology ME-USitE appears. In Section 3, the necessity of a tool of support to the evaluation is established. Section 4 describes the developed prototype. Section 5 shows the results of the study to test the methodology. Finally, in Section 6 the conclusions appear.

2 ME-USitE methodology

ME-USitE (Alva, 2005) is a methodology proposed for measure and to evaluate the usability of educative websites. The showed approach tries to complement the evaluation from the perspective of the user, being used the method of investigation and from the perspective of the expert, using inspection methods. The principal goals that Me-USitE persecutes are the following ones:

- *Evaluation multiple*: it is framed in two main evaluations: evaluation of experts and evaluation of users. The evaluation of experts, involves one or more expert making a global inspection of the application (Nielsen and Mack, 1994; Nielsen, 2002) in the different stages from the development cycle (design, development, test). The inspection can be made selected or combining the methods of inspection more suitable (heuristic, standard, guides or consistency) (Nielsen, 1994; Mack and Nielsen, 1994) to the type of site to evaluate and the experience and knowledge of the specialist. The evaluation of users, on the other hand, is an investigation evaluation in which users real – while they make tasks in a website – they discover the problems that make difficult to them to reach their objectives. This evaluation involves the use of representative end users of the objective audience to use the system and to complete a questionnaire that allows measuring the fulfillment of a set of requirements of usability measured through a set of predefined parameters.
- *Combination of methods*: the combination of the methods of inspection and inquiry for the evaluation of the usability allows discover problems that can be omitted by one of them and vice versa (Nielsen, 1994; Mack and Nielsen, 1994). To combine the usability inspections and the inquiry evaluation have as goal provide a process of more efficient and effective evaluation. It arises like an alternative proposal to the evaluation of traditional usability based on laboratory tests, whose achievement is expensive and often the obtained results are not the wished ones.
- *Configuration of the inspection methods*: because the experts are more familiarised with one or some methods, it seems logical that the inspection sees beneficiary if the expert can apply these methods. It is for that we considered a requirement for the methodology, the possibility that the expert can select the methods with which more is familiarised. Really, this will be translated in an enrichment of the evaluation process.

- *Consideration of the audience*: our methodology persecutes to evaluate the usability degree that offers the educative website to its users, considering each site is designed and developed for a specific audience and that this one can differ in level of experience due exactly to the dispersed nature of media of transmission of used knowledge (internet).
- *Extension of the evaluation to the phases of analysis and design*: most of existing methodologies fix the evaluation process to the stage of test or when the product is in use. This carries two main problems: the development of a few usable products that does not cover the necessities of the user and an increase in the cost. It is why the methodology can be applied for the evaluation from the earliest stages of the development cycle of the product.
- *Configuration of the measurement parameters*: considering that the usability requirements defer from a type of website to another one, is necessary to establish the parameters of evaluation more suitable to the type of site and type of users. Since, though the criteria of usability considered are the same, these will not affect of the same way the global usability of a site or other one, for what it is necessary to establish the most suitable parameters. Therefore, a general list of parameters of evaluation of usability in a hierarchic structure of three levels of parameters has been defined (criteria, metric and attributes), of which they are selected most suitable to the audience of the site.
- Application of a score model.

There is defined a appropriate aggregation model which allows to obtain the global score of the usability of the site, This computation is based in the parameters defined in the hierarchic structure before mentioned, starting from the lowest level parameters (attributes) up to the upper level ones (criteria).

2.1 *Evaluation model*

The process of evaluation of ME-USitE is based on a general set of measurement parameters that evaluate the usability requirements that the site must fulfill. Of these parameters selected the most suitable, regarding to the audience for which the application has been designed. On the other hand, it is necessary that categorisation of these users in groups of profiles (based on those qualities that reflect a behaviour of the user with the educative environment) that they allow to observe and to analyse the similarities within a same group and/or the differences with the others.

- *Evaluation parameters*: the process of evaluation of ME-USitE is based on a general set of measurement parameters that evaluate the usability requirements that the site must fulfill. The most suitable parameters are selected according to the audience for which the application has been designed (children, young, adult, elderly). On the other hand, considering that the educative applications, web must face an ample range of experiences and expectations from the users, it is necessary to categorise these users in groups of profiles (based on those qualities that reflect a behaviour of the user with the educative environment) which will allow us to observe and analyse

the similarities within a same group and/or the differences with the others. In this respect that have been defined, criteria, metric and attribute. To categorise the user in profiles, a questionnaire has been employed to gather the information of the users. In order to obtain the user profile using on the data obtained in the questionnaire, the next equation has been applied.

$$PU = \sum_{i=1}^4 \alpha_i \cdot EC_i$$

Where:

PU = user profile

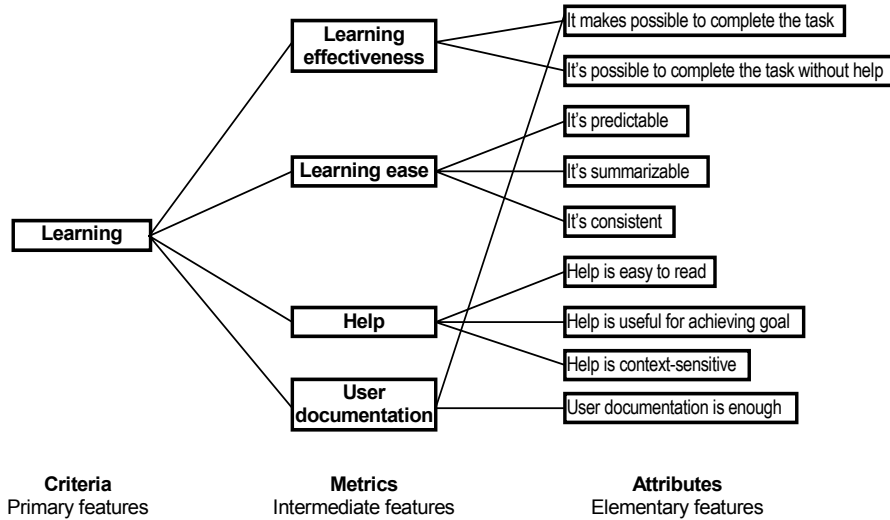
EC_i = element of profile categorisation, referred to each of the questions included in the questionnaire of information of user, according to questionnaire, four categorisation elements are considered:

- education of the user
- quantity of daily hours dedicated by the user to the use of the computer
- type of activity that the user realises with the computer
- type of software that the user has used last six months.

α_i = assigned weight to EC_i .

- *Requirements tree*: in order to carry out the evaluation process from the user perspective, an evaluation structure called requirements tree has considered itself advisable to define. The components of evaluation of this tree are defined on the basis of a hierarchic model that includes three levels: criteria (parameters of first level), metric (parameters of second level) and attributes (parameters of third level). The main purpose when designing the tree of requirements (Figure 1) is to decompose the parameters of evaluation (criteria) in elements simpler than they facilitate the assignment of weights and the calculation of scores and for that reason must be smallest possible but trying to include the most important aspects in function to the objective audience.
- *Measurement model*: once the parameters have been established, it is necessary to associate them a quantitative value, according to the measurement model proposed by the authors. The model used is based on the evaluation model cost/benefits logic scoring of preference (LSP) proposed by Dujmovic et al. (Ziljlstra, 1993), that was developed for the comparison and selection of alternatives of systems of hardware or complex software and that uses aggregation functions to reflect the preference of the evaluator users. This approach applies two criteria: analysis of the preference in the quality and analysis of the cost. Nevertheless, since we are interested in determining the level of usability of a site, we have not considered in our model the cost analysis. This model has been used (Nakwichian and Sunetnanta, 2003) for the valuation of quality of websites on the user based.

Figure 1 Hierarchical structure of parameters in *learning tree*



2.2 Evaluation process

The parameters defined in the hierarchic structure will be the base for the calculation of the global score of the site, for it will become use of a set of functions of aggregation appropriately established that are in charge to calculate the score starting off of the parameters of lower level (the attributes) until those of more high level (criteria). These functions take like entrances the obtained elementary scores from the following way:

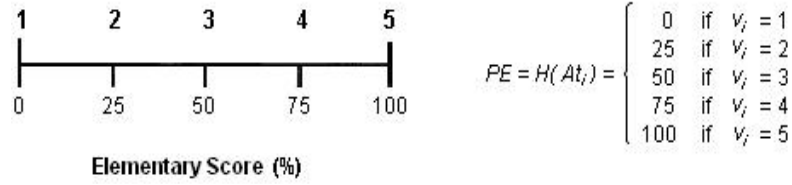
- The user assigns a numerical value to each attribute (elementary parameter) $At_i, (i = 1, \dots, n) : v_i, \forall i = 1, \dots, n$ (n = number of attributes of the sub-tree), in a range of possible values in a scale of valuation 1–5 (agreement-disagreement). The value assigned by the user is standardised, by means of a function of transformation PE whereby it becomes an elementary score or of attribute on a scale from 0–100, that it indicates the degree of conformity of the usability parameter with respect to the requirements established for the application domain that is being evaluated.

$$1 \leq v_i \leq 5,$$

where:

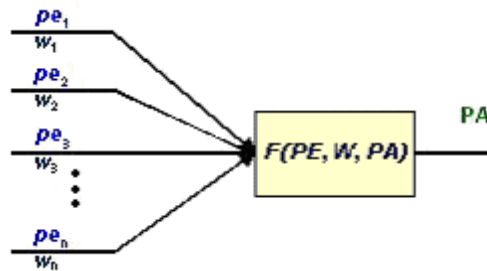
$$v_i = \begin{cases} 1 = & \text{Disagreement strong} \\ 2 = & \text{Disagreement} \\ 3 = & \text{Neutral} \\ 4 = & \text{Agreement} \\ 5 = & \text{Agreement strong} \end{cases}$$

- The scores aggregated obtained in a level, will become as well in new entrances for the functions of aggregation of the following level. Finally, the process will be repeated so that from those functions the global assessment of the site is obtained.



In Figure 2, the application of the aggregation function (F) a group of elementary scores within a sub-tree is showing.

Figure 2 Model of aggregation of preference of sub-tree k (see online version for colours)



The equation defined to obtain the aggregate score for every level of the structure of evaluation and the global score of the site will be:

$$PA = \left(w_1 pe_1^r + w_2 pe_2^r + \dots + w_n pe_n^{1/r} \right)$$

where:

w_i assigned weight to each one of the evaluation elements to determine its importance, being:

$$\sum_{i=1}^n w_i = 1$$

pe elementary score obtained in the previous level

r factor associated with the type of parameters to evaluating (essential, desirable or optional) and the number of them.

2.3 Criteria of acceptance of the level of usability

The criteria for the acceptability of the level of usability of an educational website are based on the proposed ones for the ISO standard 14598 (ISO99) in it that establishes three regions in a range from 0 to 100 %: satisfactory, acceptable and unsatisfactory.

The determination of the range of the acceptable region established with a score between 40 to 60 %, is based on the criteria of reliability established in the analysis of results realised in the use of questionnaires like SUMMI (VEE98) and MUMS (VEE03)

in which a scale from 0 to 100 with an average of 50 and one standard deviation of 10 % is used.

3 Need of a user evaluation tool

If we considered that the evaluation process consumes long time, since involves by a side the application of questionnaires to users and the use of lists of verification on the part of the experts and on the other hand implies the information processing thus obtained to achieve concrete results, seems necessary to use automated tools. These tools will as much give to support to the process of evaluation for the obtaining of data as the treatment and processing of such.

Table 1 Comparison of usability evaluation tools

<i>Tool</i>	<i>Methods combination (inquiry/ inspection)</i>	<i>Applied technical (questionnaire/ check list)</i>	<i>Stage of the development (design, development, test)</i>	<i>In focus (user and expert)</i>	<i>User profile</i>	<i>Audience</i>
WAMMI	X	√	√	X	X	X
ISOMETRIC	X	√	√	X	X	X
MUMS	X	√	X	X	X	X
PROKUS	√	√	√	X	X	X
QUIS	X	√	X	X	X	X
SUMI	X	√	X	X	X	X
DRUM	X	X	√	X	X	X
SMEQ	X	X	X	X	X	X
TLX	X	X	X	X	X	X

Source: Alva et al., 2003.

In order to cover these needs it is necessary to look for a tool that helps to automate the process proposed by ME-USitE. For that reason, a study of the existing tools of support for the usability evaluation has been made (Alva, 2005). The examined tools are: WAMMI (Kirakowski et al., 1998), ISOMETRIC (Gedica et al., 1999), SUMMI (Bevan, 1995), PROKUS (Zülch and Stowasser, 2000), DRUM (Macleod and Rengger, 1993) and SMEQ (<http://ed.sjtu.edu.cn/ranking.htm>). We have reviewed each one of the mentioned tools previously, in order to analyse if some of them adapts to the methods and techniques proposed in ME-USitE. As we can see in the table, the tools have been designed in basis of a perspective (user), none of them takes into account aspects like the profile from the evaluator user or the level of audience of this user, aspects that the methodology considers fundamental in the evaluation process.

3.1 Desirable characteristics of the evaluation tool

The tool must be able to offer a support for the evaluation and the processing of the data collected according to the proposed methodology. Therefore, the goal pursued in the

development of a tool of this type will be grouped around two fundamental aspects: the evaluation and the processing.

Support for the evaluation: to achieve this target it must provide the following characteristics:

- *Support for different types of users:* the tool must distinguish from different types of users: evaluator user, recorder user, manager user.
- *Control of the evaluation sites:* the tool will allow the recorder user to register his/her site for the evaluation. These users must provide some information such as: URL (site address), name of the site, information of the person in charge, mail of contact, type of educational site, audience, critical tasks to evaluate, use scenes etc.
- *Determination of the level and profile of the evaluator user:* the tool must display a questionnaire for the evaluator user to provide both general and specific information about them. This information will be used to classify them within a definite audience (children, young, adult, elderly). In addition to this, it must allow a mechanism to establish the profile of the above mentioned user (inexperienced, intermediate and advanced).
- *Storage of the tree of requirements:* the tool will have to store both the tree of requirements of usability (criteria, metric and attributes) established for each level of user and the questionnaire defined for each case. This way, during the evaluation, and once the user is identified; the tool will be able to identify the attributes that this user will evaluate.
- *Selection of the site to evaluate:* the tool has to provide the user with the list of sites to evaluate so that he can select one of them in accordance with his/her preferences or interests. Nevertheless, the tool will also have to allow the user to propose suggest the particular site that he/she wants to evaluate (taking into account that the user can be the apprentice, professor, tutor or visitor interested in the area).
- *Selection of tasks:* the tool must allow the selection of one or more tasks for the evaluation and a possible use setting (in case the user makes the selection of a site). When the evaluator user proposes a site to evaluate, the tool has to allow the description of the tasks and settings to evaluate.
- *Configuration of the evaluation:* the tool must offer the user flexibility of evaluation, that is to say, it must allow the user to decide if he/she wants to carry out a complete evaluation (to evaluate all the parameters of the pre-established tree of requisites at his level) or a partial evaluation selecting the parameters of the tree that he/she wants to evaluate (that is to say, criteria, metric and specific attributes). This partial evaluation is allowed in order to establish the aspects of the evaluation that are of major interest for certain evaluator users.
- *Application of the questionnaire of evaluation:* as soon as the site to evaluate and the critical task(s) are selected, the tool will have to identify the questionnaire to apply in accordance with the tree of requisites which is suitable for the evaluator user. The tool will have to allow the user to gain access to the elected site and to carry out the task(s) programmed to answer the questionnaire that should allow the evaluation of the chosen attributes.

- *Update of the requirements tree*: the tool should allow the incorporation and/or the change of the parameters in the requirements tree, especially due to the experimental character of this work.

Support for the measurement and processing: the tool must provide the following characteristics:

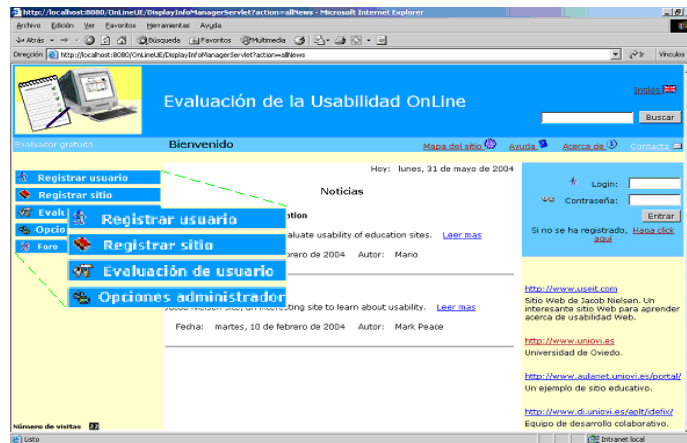
- *Calculation of the punctuation of attributes*: the answers for the corresponding attributes in the requirements tree that were provided in the questionnaire of evaluation will have to be assigned and stored (measurable parameters directly) to be used in the calculation of the usability score. In addition to this, the tool will have to support the presentation of this information (to tabulate, graph and even make possible the option of printing or download).
- *Calculation of the usability score*: from the punctuations of the attributes, the tool will allow/permit to calculate the punctuation of every level of the requirements tree (metric and criteria). This calculation will have to be carried out as many times as the users take part in the evaluation. The tool will also have to store these results and provide its presentation in different formats.
- *Obtaining of the global and entire score*: based on previous results, the tool will have to calculate the entire usability score of the site.
- *Obtaining of partial scores*: the tool has to allow specific partial results such as total punctuations for the profile both better and worse parameters valued globally or within each profile and partial results of metric and criteria within the profile.
- *Obtaining of the score for user's level*: if the site to evaluate has users of more than one level (children, young, etc.) as the target audience, the tool will have to classify the results obtained in accordance with the user's level, facilitating the presentation of results in different formats and in a wide or abridged way.
- *Obtaining of special results*: in addition and based on the information that could be obtained from the user's questionnaire, the tool must help to obtain special results on the usability evaluation, such as, for instance, the punctuation for genre (to determine the parameters which are better and worse valued for each genre) and punctuation for the hand used in the handling of the mouse (to analyse the level of accessibility perceived by right-handed and left-handed users).

4 Proposed tool for ME-USitE

Finally and to cover the needs previously exposed (Section 3), it has been decided to construct the first prototype at present in execution. The characteristics of this prototype that coincide basically with the needs previously explained are provided across three modules: module of registration, module of evaluation and module for the basic management of the application.

Figure 4 shows the interface of the tool (PES04), where three modules are/appear reflected.

Figure 4 Tool to support the evaluation of the user of educational web sites (see online version for colours)



- *Registration module*: the designed prototype allows identifying three types of users: evaluator, recorder of the educational website and manager. To cover the needs mentioned in Section 3.2, this module includes: the register of the site to evaluate and the register of the evaluator user.
- *Evaluation module*: this module takes charge of the necessary management to realise the evaluation on the part of the user, as well as the presentation of the results. It includes the following functionalities: determination of the level and user's profile, automatic configuration of the usability requirements tree, selection of the website to evaluate, configuration of the evaluation (selection of tasks, application of the questionnaire of evaluation and calculation of the scores).
- *Module for the basic management of the application*: for the suitable functioning of the prototype, it is necessary information such as: authorised users, websites to evaluate, requirements tree, etc. For the experimental character of the prototype, it is also necessary to be able to adapt and to update the above mentioned information. Normally this task will be carrying out by the manager user: control of access, control of sites and management of users, update of the requirements tree, help and documentation.

5 Methodology

A study was conducted to contrast the obtained results applying our proposal of evaluation with regard to a similar one, such as Olsina's proposal (Paolini and Di Blais, 2002; Preece, 1993) that proposes an evaluation of the quality of websites and in among other aspects, it considers the evaluation of the usability. For the evaluation they were considered to be five better Spanish universities of agreement to the world ranking (Academic Ranking of World Universities, 2008): University of Valencia, Autonomous University of Barcelona, Complutensian University of Madrid, Autonomous University of Madrid and University of Barcelona.

For the study one relied on 68 participants, whose ages were ranging between 18 and 28 years (with an average of age of 22.5), distributed of a following way: 28 students of the School of Teaching, 16 of the School of Computer Science and 24 of the Máster in Web Engineering, all students of the University of Oviedo.

The participants were invited to evaluate the sites of the selected universities using both approaches: the criteria of usability established in the methodology of evaluation of the quality website QEM of Olsina et al. (1999) and the questionnaire of evaluation proposed by our methodology. The questions included in our questionnaire were obtained as result of evaluating the profile of the evaluator user across a previous questionnaire completed by this one for such a purpose (configuration supported by the tool).

5.1 *The data analysis and results*

Once the experiment has been made, the final data of the evaluations was collected and processed by the proposed tool for ME-USitE. In the first case (evaluation applying Olsina's criteria), the process of calculation was realised manually using his model of valuation, for our proposal, this was not necessary since it was carried out for the support tool. Finally, it showed the results of evaluation for the different universities:

Table 2 Summary of evaluation of the evaluated universities

	<i>Website QEM</i>	<i>ME-USitE</i>
University of Valencia	0.65	0.74
Autonomous University of Barcelona	0.75	0.88
Complutensian University of Madrid	0.6	0.69
Autonomous University of Madrid	0.55	0.6
University of Barcelona	0.38	0.49

According to the criteria of acceptance seen in Section 2.3 and used so much in our proposal as in that of Olsina, the level of acceptability of the usability of the evaluated sites was the following one:

Table 3 Level of acceptability

	<i>Olsina</i>	<i>ME-USitE</i>
University of Valencia	Good	Good
Autonomous University of Barcelona	Good	Good
Complutensian University of Madrid	Acceptable	Good
Autonomous University of Madrid	Acceptable	Acceptable
University of Barcelona	Poor	Acceptable

In addition, our proposal has the added one and is that it allows not to obtain a global score of usability as Olsina's case, but in addition, it offers partial results in the levels of criteria and metrics, evaluations for type of user's profile (inexperienced, intermediate, advanced), as well as to obtain the degree of correlation between the assessment obtained by the user's profiles different. In addition, offering score by sex, manipulation of the mouse (right, left-handed), groups of user (children, young, adults and elders).

Table 4 shows the results of the parameters of higher level (criteria), obtained for every evaluated university, applying our offer.

Table 4 Summary of evaluation by criteria

	<i>University of Valencia</i>	<i>Autonomous University of Barcelona</i>	<i>Complutensian University of Madrid</i>	<i>Autonomous University of Madrid</i>	<i>University of Barcelona</i>
Learning	0.83	0.84	0.83	0.70	0.49
Operativity	0.85	0.90	0.75	0.72	0.6
Content	0.91	0.95	0.83	0.80	0.82
Attractivity	0.75	0.80	0.75	0.33	0.71
Satisfaction	0.76	0.75	0.83	0.63	0.46
Communication	0.50	0.52	0.51	0.46	0.57

6 Conclusions

With the main goal of a methodology for assessing and measuring educational websites usability as ME-USitE, arises the need of a tool to systematise the process of evaluation. Nevertheless and due to the approach proposed in the methodology, focused on the audience and user's profile, of the existing tools, there has not been one that satisfies the needs raised in the Section 3.2. So it was necessary to propose the development of a tool adapted to our proposal.

In this article, we described the design of a tool that covers the functionalities needed to contribute in the detection of problems and mistakes of usability of educational websites from the user's perspective in his environment of real work. The results obtained across the tool enables the capture of decisions for website usability improvement, since they provide a score that determines if the level of usability reached by the site covers the requisites outlined for the proposed audience and the user's profile taking part in the evaluation.

For the development of the tool it was proposed that this one fulfills two principal functionalities: support to the evaluation process, where this first functionality aspects are covered as the identification of the profile of the user and the determination of the group of the audience to which the evaluator belongs in order to identify the corresponding requirements tree as well as the questionnaire to apply and the support to the query of the gathering data. As soon as the evaluation was carried out, the obtained data will have to be processed in order to obtain results of the usability of the site. In this sense, the tool calculates the scores added for attributes, metric and criteria being based on the evaluations given by the users and it show the obtained results of a partial or complete site, profile, genre, criterion, metrics, attribute or hand of the handling of the mouse.

Of the obtained results, we have found differences between applying the approach of Olsina and our approach, according to Tables 2 and 3. Olsina considers the usability as an aspect of the quality of a product together with the efficiency, reliability and functionality. Including in the evaluation of the usability aspects of understandability, feedback and aesthetic, whereas our proposal which is focused on the usability of educational domain sites includes in addition educational aspects of communication and learning methods. For that reasons, we think that in the case of the University of Barcelona, applying Olsina's approach, this site fulfills 38% of the requirements of usability with which it is located in the region of dissatisfaction, with a level of poor usability, whereas applying our approach this percentage is 49%, being located in the

satisfactory region, though with an acceptable level. Similar case happens for the Complutensian University of Madrid, which though in both approaches is satisfactory, applying Olsina would be acceptable whereas for us would be good.

References

- Academic Ranking of World Universities*, available at: <http://ed.sjtu.edu.cn/ranking.htm>.
- Alva, M., Martinez, A., Cueva, J., Sagástegui, H. and Lopez, B. (2003) 'Comparison of methods and existing tools for measurement of usability in the web', *ICWE 03: Third International Conference on Web Engineering 2003*, Spain, LNCS/LNAI Proceedings, Springer Verlag.
- Alva, O., María, E. (2005) 'Metodología de evaluación y medición de la usabilidad de sitios web educativos: ME-USitE', Tesis inédita, Universidad de Oviedo, España.
- Bevan, N. (1995) 'Measuring usability as quality of use', *Software Quality Journal*, Vol. 4, pp.115–150, NPL Usability Services, Teddington–Middx, TW110LW.
- Carvalho, A. (2002) 'Usability testing of educational software methods, techniques and evaluators', *3 Simposio Internacional de Informática Educativa*, Portugal.
- Celik, H. and Ipcioglu, I. (2007) 'Gender differences in the acceptance of information and communication technologies: the case of internet usage', *International Journal of Knowledge and Learning*, Vol. 3, No. 6, pp.576–591.
- Gedica, G., Hamburg, K. and Duentsch, I. (1999) 'The IsoMetrics usability inventory: an operationalization of ISO 9241-10 supporting summative and formative evaluation of software systems', *Behaviour and Information Technology*, Vol. 18, No. 3, pp151–164, Ireland.
<http://www.nlc-bnc.ca/9/1/p1-260-e.html>
<http://www.ucc.ie/hfrg/questionnaires/mumms/info.html>
- Human Computer Interaction Lab/University of Maryland (1998) 'QUIS-questionnaire for user interaction satisfaction'.
- Human Factors Research Group, *MUMMS: Measuring the Usability of Multi-media Systems*, University College Cork, Ireland.
- Kirakowski, J., Claridge, N. and Whitehand, R. (1998) 'Human centered measures of success in web site design', *Human Factors and the Web Workshop*, Basking Ridge, NJ.
- Mack, R.L. and Nielsen, J. (Eds.) (1994) 'Executive summary', *Usability Inspection Methods*, John Wiley & Sons, New York.
- Macleod, M. and Rengger, R. (1993) 'The development of DRUM: a software tool for video-assisted usability evaluation', *In People and Computers VII*, Cambridge University Press.
- Mineta, N. (2000) *Americans in the Information Age – Falling Through the Net*, U.S. Department of Commerce.
- Murray, G. and Constanzo, T. (1999) *Usability and the Web: An Overview*, Canada.
- Nakwichian, S. and Sunetnanta, T. (2003) *User-centric Web Quality Assessment Model*, Mahidol University Computing Center, Bangkok 10400, Thailand. Revisado en Diciembre.
- Nielsen J. (1994) 'Heuristic evaluation', in J. Nielsen and R.L. Mack (Eds.): *Usability Inspection Methods*, John Wiley & Sons, USA.
- Nielsen, J. (2001) 'The web usage paradox: why do people use something this bad?', *Alertbox*, 09/081998 (current 03/01/2001), available at: www.useit.com/alertbox/980809.html.
- Nielsen, J. (2002) *Usability Inspection Methods*, SunSoft, Mountain View, CA., Copyright ACM.
- Nielsen, J. (2005) 'Usability of websites for teenagers', *Alertbox*, 31/01/2005.
- Nielsen, J. and Mack, R. (1994) *Usability Inspection Methods*, John Wiley & Sons, New York, ISBN0-471-01877-5.

- Nielsen, J. and Molich, R. (1990) 'Heuristic evaluation of user interfaces', *Proceedings of the CHI'90*, New York, pp.349–256, ACM 0-89791-345 0/90/0004-0249.
- Olsina, L., Godoy, D., Lafuente, G. and Rossi, G. (1999) 'Assessing the quality of academic web sites', *New Review Hypermedia Multimedia Journal*, Vol. 5, pp.81–103, Taylor Graham Publishers, UK, ISSN 13614568.
- Paolini, P. and Di Blas, N. (2002) 'Automatically evaluating the usability of web sites', *CHI 2002 Workshop*, Italy.
- Preece, J. (1993) *A Guide to Usability: Human Factors in Computing*, Addison Wesley, The Open University.
- Shneiderman, B. (1998) *Designing the User Interface*, Addison-Wesley, Reading, MA.
- Su, S., Dujimovic, J., Batory, B., Navathe, S. and Elnicki, R. (1987) 'A cost-benefit decision model: analysis, comparison and selection of data management systems', Vol. 12, No. 3, pp.472,320.
- Tromp, J. and Benford, S. (1997) 'Methodology of usability evaluation for collaborative virtual environments', *Proceedings of 4th UK VR SIG Conference*, Bristol.
- Zijlstra, F. (1993) *Efficiency in Work Behaviour: A Design Approach for Modern Tools*, Delft University Press, Delft.
- Zülch, G. and Stowasser, S. (2000) *Usability Evaluation of User Interfaces with the Computer Aided – Evaluation Tool – PROKUS*, Institut für Arbeitswissenschaft und Betriebsorganisation (ifab), Universität Karlsruhe (TH), Kaiserstraße 12, 76128 MMI-Interaktiv, Nr. 3, Zülch & Stowasser, Germany, ISSN1439-7854.