

# Project and Development of ErgoCoIn Version 2.0

Marcelo Morandini, Roberto Leite de Moraes Rodrigues,  
Marcus Vinicius Cerrato, and Marcos Lordello Chaim

School of Arts, Science and Humanities, State University of Sao Paulo, Av. Arlindo Bettio,  
1000, Ermelino Matarazzo, Sao Paulo, SP, Brazil, 03828-000  
{m.morandini, roberto.rodrigues, mcerrato, chaim}@usp.br

**Abstract.** Usability evaluation (UE) methods may have several purposes: identify and diagnose usability problems; evaluate the implementation, comparing the one under evaluation with other systems, and also perform acceptance test. These methods are usually expensive, time-consuming and complex. In this context, the ErgoCoIn 1.0 has been used as an accessible system that can be used by several people all over the world, even if they are not usability experts or evaluators. This paper aims to present a new version of this environment considering that its basic objective is to propose a higher quality for the questions. Additionally, a new presentation questionnaire window small enough to be used both in web browsers and mobile devices is proposed.

**Keywords:** Usability Evaluation, Checklist, Automation.

## 1 Introduction

Software developers and Human-Computer Interaction (HCI) designers always deal with a common problem: evaluate the usability of the systems they developed. Specifically, concerning the internet based systems, that are usually under constant evolution, their features can be analyzed as the former *web sites* [4].

Usability, as pointed by the ISO/IEC 9241, can be defined as: the *effectiveness, efficiency, and satisfaction* with which users achieve specified goals in particular environments. Effectiveness can be defined as the accuracy and completeness with which specified users can achieve specified goals in particular environments. Efficiency as the resources expended in relation to the accuracy and completeness of goals achieved. And, finally, user satisfaction as the comfort and acceptability of the work system to its users and other people affected by its use [2].

Specifically, some web sites usability evaluation techniques [1; 2; 3; 6; 7; 8; 10] use previous evaluations approaches that produce qualitative reports, which might lead to subjectivity problems. This paper presents an objective evaluation approach that aims to quantify the efficiency, efficacy and user satisfaction while interacting with a web site to perform pre-defined tasks, and this approach is based on the server log files registers. We wish to confirm that it is possible to go further in the log files analysis in order to present usability studies. Using the log files data mining, it is possible to obtain taxes and metrics that quantify the usability of some tasks, producing trustworthy, objective, fast and low cost evaluations.

A basic demand of any website is that its web application should work to fulfill its intended purpose – this is addressed by research into Web Engineering methods [1, 13, 14], models and tools. However, there is also the equally important demand that the web application must be usable by the visitors of the website. Usability research includes work on sets of guidelines which help to improve website usability. The ErgoCoIn environment [12] is an approach designed to give support to non-ergonomic expert evaluators in such a way they could conduct contextual, objective and systematic usability evaluations of web sites and their pages. It postulates the identification of the context of use aspects, comprising: users, tasks and environment attributes concerning the web site under evaluation. Content of the interviews/questionnaires [11] and the other environment data gathering activities is based on information demand underlined by the base of ergonomic questions initially compiled in this project. Such a strategy allows for proceeding very objective ergonomic inspections. Only pertinent information are proposed to inspectors in the contextual analysis and only applicable questions are presented to them in the web site inspection. Objectiveness is obtained also by inspections questions that are linked to a logic construction based on ergonomic criteria [9] that define their relative importance.

## 2 Web Site's Usability Evaluation

The usability of a product is affected not only by the features of the product itself, but also by the characteristics of the users, the tasks they are carrying out, and the technical, organizational and physical environment in which the product is used.

The main quality in an User Interface (UI) is its usability, which establishes how much the systems projected are easy to use and learn [1, 6]. Usability Evaluation (UE) methods have several purposes, such as: to aid the Web UI project to reach the user needs; to identify and diagnosis problems; and to perform a cognitive usability evaluation (comparing with other systems, and also for the acceptance test). The collected data during these UEs can be qualitative (descriptions that characterize the usability) or quantitative (measured that quantify the usability). Some interactive systems UE methods found in literature [1; 4; 5, 6, 10] approaches UI qualitative aspects; however they are not enough objective to quantify the efficiency, effectiveness and user satisfaction.

Therefore, UE is an important part of the overall user interface design process, which consists of iterative cycles of designing, prototyping, and evaluating [8]. Usability evaluation is itself a process that entails many activities depending on the method employed [13]. Common activities include: *Capture* collecting usability data, such as task completion time, errors, guideline violations, and subjective ratings; *Analysis* interpreting usability data to identify usability problems in the interface; and *Critique*: suggesting solutions or improvements to mitigate problems [4].

A wide range of usability evaluation techniques have been proposed. Some evaluation techniques, such as formal user testing, can only be applied after the interface design or prototype has been implemented. Others, such as heuristic evaluation, can be applied in the early stages of design. Each technique has its own requirements, and generally different techniques can uncover different usability problems. Usability

findings can vary widely when different evaluators study the same user interface, even if they use the same evaluation technique [1,4,10].

Automation of usability evaluation has several potential advantages such as the following:

- Reducing the cost of usability evaluation.
- Increasing consistency of the errors uncovered.
- Predicting time and error costs across an entire design.
- Reducing the need for evaluation expertise among individual evaluators.
- Increasing the coverage of evaluated features.
- Enabling comparisons between alternative designs.
- Incorporating evaluation within the design phase of UI development, as opposed to being applied after implementation.

It is well known that the average quality of websites is poor, “lack of navigability” being the main cause of user dissatisfaction [10]. On the one hand web technologies evolve extremely fast, enabling sophisticated tools to be deployed and complex interactions to take place. Secondly, the life cycle of a website is also extremely fast: maintenance of a website is performed at a rate that is higher than that of other software products because of market pressure and lack of distribution barriers. In addition, often the scope of maintenance becomes so wide that a complete redesign takes place.

On the other hand, the quality of a website is rooted on its usability, which usually results from the adoption of user-centered development and evaluation approaches [1, 10, 12, 13, 14]. Usability testing is thus a necessary and repeated step during the life-cycle of a website. To test usability of a website a developer can adopt two kinds of methods: usability inspection methods (e.g. heuristic evaluation [1, 5, 13] or user testing [9]). Heuristic evaluation is based on a pool of experts that inspect and use a (part of a) website and identify usability problems that they assume will affect end users. With *user testing*, a sample of the user population of the website is selected and is asked to use (part of the) website and report things that they think did not work or are not appropriate.

Even though the cost (in terms of time and effort) of both methods is not particularly high, and their application improves the website quality and reduce the overall development cost, they are not systematically performed at detailed levels on every different part of a website after each maintenance or development step.

Specifically, the checklists can be applied by design staff, not necessarily experts in usability evaluation, such as: programmers and users. The checklists are extremely cheap to be proposed and only need some few users to be effective. They can also be virtually proposed at any time in the development life cycle, i.e., from the very initial HCI prototype until a concluded and under use web site [11, 14].

### 3 The ErgoCoIn Version 1.0

Web sites development became accessible to a great amount of designers, not just the experts on usability or ergonomics fields. So, if the developments are always faster, the products developed can present a lack of quality [15]. According to this, we have

defined an approach for the ErgoCoIn environment considering that there is the need that non-expert designers should be presented with a minimal quantity of information concerning a strategy for their designs as the context of use must be well defined and understood. A second consideration is that web sites are often designed through a fast and low cost design processes supported by tools which lead designers to carry out numerous and, sometimes obvious, ergonomic flaws.

ErgoCoIn is a checklist-based environment initially aimed at supporting objective ergonomic inspections of e-commerce web site and pages [16]. The ErgoCoIn features automatic inquiring services to identify context of use aspects (users and environment attributes) and to recognize web page components. Consequently, it is able to propose to inspectors only questions applied to the specific task context of use and to the associated web page components [12]. The ErgoCoIn validation phase was based on case studies composed by repetitive cycles of inspections to be carried out by a group of inspectors. The idea is to verify objectiveness and systemic attributes like results repeatability and reproducibility at the end of each cycle. In the last cycle, problematic questions can be deleted. These case studies will also be used for the analysis of the inspection tasks and for the specification of the most fitted user interface for support this task. ErgoCoIn logic architecture can be filled in by inspection questions and criteria concerning the different kinds of web sites application domains. Initially, the original knowledge content is supposed to be specific to e-commerce, due to its importance and also to the high number of web sites in this domain that can be used to validate this environment.

Aiming the ErgoCoIn's definition, two evaluation techniques have been selected: Inspection Methods and Questionnaires. Inspection Methods do not require the direct participation of users (as it is easier and less costly) [9, 12]. They regroup a set of approaches that use the evaluators' judgment, being experts or not in usability. Even if these various methods may have different goals, they generally aim to detect the interface features that can lead to inefficiencies or higher workload for the users. The inspection methods distinguish themselves by the way the evaluators' judgment is performed and by the evaluation criteria that are the basis for the judgments. Specifically, we selected inspection techniques that are strongly supported by heuristics or even better, by individual recommendations.

The questionnaires allow not only to collect subjective data related to the users' actions and comments about their satisfaction from web sites, but also about the achievement of factual elements concerned to the system under evaluation [11]. The interviews/questionnaires with users can be used to identify features of the context of use, the usability difficulties, and collect local meaningful users' comments. They are even considered as a way to establish comparisons between the scenarios prescribed by the designer and the "real" ones, as described by the user.

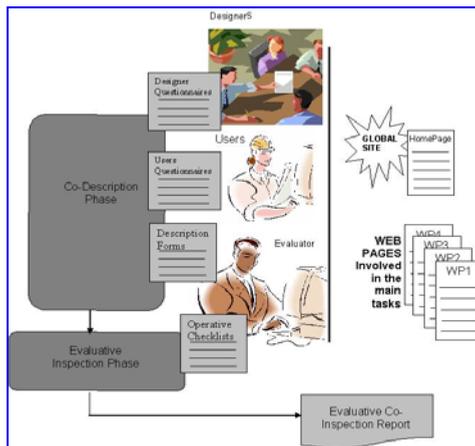
The design of the usability evaluation technique underlying the ErgoCoIn has been motivated by four considerations: (1) that web sites development became accessible (through easily available design tools) to a large spectrum of "designers", not necessarily highly skilled in computer science or in ergonomics; (2) web sites are often designed along a fast and low cost design process supported by non expensive tools which lead designers to carry out numerous and sometimes obvious ergonomic flaws; (3) the high cost of usability evaluations using the "traditional" methods; and (4) the subjectivity that is a common problem in any usability evaluation.

The constraints listed above lead to the main characteristic of the intended method: an inspection-based evaluation with directly accessible web-based support mechanisms (strategies, recommendations, etc.), and information on context of use coming from users and designers. Several parameters must be set when selecting components for an evaluation method.

Therefore, it was defined that all this knowledge should be directly collected from the users and also from the designers using questionnaires or interviews for this purpose. So, due to this definition, the approach only can be applied to real web sites that are being used regularly and it is necessary to have the designers and some users available for the interviews or, at least, able to answer some questionnaires [11].

ErgoCoIn combines the Inspection Method and Questionnaires Evaluation techniques in an approach able to allow rapid and context focused ergonomic inspections [4, 5, 12]. The inspection component resulting from examining a large collection of ergonomic recommendations later completed with other data collected from different studies to elaborate checklists for the ergonomic characteristics applicable on web sites [8, 9]. These recommendations were formulated as questions and associated with both an ergonomic criterion that allows defining a system of relative importance between questions. Interviews/questionnaires and guidance for collecting data from users and designers were defined from analyzing the information demands in each question we elaborated. ErgoCoIn approach's is divided into two main phases: web site's Contextual Analysis and its Evaluative Inspection, as presented in Figure 1.

The ErgoCoIn's Contextual Analysis phase collects all information related to the web site operational contexts that are useful for the usability evaluation process. This phase consists of a site description process and interviews of the users and designers. Also, ErgoCoIn supports and automatizes these activities. A HTML component recognizing tool identifies the existence of specific web site user interface components and aspects and organizes them according to two categories of descriptions: the global web site and the individual web pages associated with the main tasks accomplishment [12].



**Fig. 1.** The ErgoCoIn approach framework [12]

The second phase of the method is formed exclusively by inspection activities. ErgoCoIn starts the process performing an automatic analytical evaluation based on the comparison between user's and designer's information concerning the intended and the real context of use features. The system points out to existence of designer's misconceptions about user's features, and indicates the web site aspects to verify or reformulate in consequence.

At that, the system assembles checklists concerning the overall site and the specific web pages features related to task scenarios. These checklists can be considered as "objective" ones, once they will propose only the site components applicable questions arranged according to their levels of importance [12]. Applicability decisions result from processing the site description stored in the context of use database. Priority decisions results from ranking the Ergonomic Criteria [9] according to usage context features. A default Ergonomic Criteria ranking is suggested, but it can be modified by the evaluators, according to the characteristics of the current web site context of use. In fact, the original importance structure was proposed with a general Business to Consumer usage context in mind, in which non professionals users operate sites of virtual stores from their home environments aiming to buy simple products in a relatively low frequent basis. In such a situation the Guidance criterion should be considered before the Workload criterion. Anyway, ErgoCoin can authorize evaluators changing the importance structure at the ergonomic criteria level to accommodate different usage contexts.

## 4 The ErgoCoIn Version 2.0

This section aims to present a new version of the ErgoCoIn environment considering that its basic objective is to propose a higher quality for the questions that should be applied in the questionnaires that may be proposed. This new questionnaire proposition allows the definition and realization of objective ergonomic inspections by gathering pertinent information concerning the web site's supported tasks, proposing only pertinent questions that should be presented to the inspectors that are relevant and significant and can be answered during the web site's inspection. So, among others, we are updating the following activities:

1. A presentation questionnaire window evolution that now allows the filling of the answers in a smaller and second level window. This window only has one visible question per time and has visible interactions buttons presented, such as next and previous questions and help. This new small window will always be visible to the ErgoCoIn's users and it may be possible to move it over the screen in order to allow the inspectors to interact with the web site to answer the question. Also, we are proposing the dimensions for this new window to be small enough to be used both by web browsers and mobile devices.
2. The questions that compose the questionnaire can be updated (deleted, inserted or modified) by the evaluator, even if these questions may belong to any context group or are related to any specific ergonomic criterion [9]. This aims to avoid "Not Applicable" answers for the questions.
3. Each one of the questions is now related to a specific Ergonomic Criterion. This is important for future analysis, which includes a quantitative ranking calculation based on the criteria.

As in the version 1.0, the evaluative inspections are performed by an evaluator applying the set of checklists defined in the previous phase. As mentioned before, this process constitutes an evaluative inspection as the evaluator is asked to judge the quality of very precise web site features. The questions in this method could not be answered by an intelligent or expert system once they require consideration of several factors not always at the reach of such systems. However, the level of judgment proposed by questions was defined in accordance with the level of ergonomic knowledge expected from evaluators (fairly basic usability expertise). Indeed, the questions phrases and associated support information, like justification and examples, were formulated in order to be easily understandable. Even so, this method leads to results which are subjective in nature and which may not point at all major ergonomics problems. As mentioned earlier, it is viewed as a rapid and widely usable mean to detect usability design faults, but in which human involvement is required.

Figure 2 presents the description window for the questions that compose the questionnaire to be applied. Although this window is still written in Portuguese, it is easy to realize that there are important information presented:

1. a specification of the main ergonomic criterion;
2. the presentation of the question itself;
3. a description that explains the question, including important information, such as references and links; and
4. the kind of the answer that is divided in two possibilities – Yes/No or a leveled answer, that represents the possibility of given a score as an answer, such a 0 – 10 ranking.

The screenshot shows the 'ErgoCoIn Web' interface with a navigation bar containing 'Administration', 'Contextualization', 'Settings', and 'Rating'. Below the navigation bar is a 'Questions' section containing a table with the following data:

Id	Criterion	Statement	Description	Type
<a href="#">Edit</a>   <a href="#">Delete</a> 1	Promptitude	Titles of screens, windows and dialog boxes are at the top centered or left justified?	Titles of screens, windows and dialog boxes. Titles of screens, windows and dialog boxes must be on top or centered or aligned left. Reference: Bodart and Vanderdonck [1993] p. 136 rec 4	Yes/No
<a href="#">Edit</a>   <a href="#">Delete</a> 3	Feedback	The system provides feedback for all user actions?	The system must acknowledge receipt of an immediate action from all introduced by the user, who must be able to recognize it as an apparent reaction system. Reference: Smith and Moser [1986] p. 15 rec 1.0.3	Graduated
<a href="#">Edit</a>   <a href="#">Delete</a> 9	Grouping by location	The presentation space is diagrammed in small functional areas?	Screens, windows and boxes with functional zones diagrammed. The presentation space must be diagrammed in small functional zones. POSITIVE EXAMPLE: Put small functional areas towards the perimeter of the screen. Reference: Bodart Vanderdonck [1993] p. 144 2 rec	Yes/No
<a href="#">Edit</a>   <a href="#">Delete</a> 10	Grouping by location	Groups of control buttons are arranged in column and right, or on-line and below the objects to which they are associated?	Every group of command buttons on the same logical set of data must be prepared, be on the line below the object on which they are understood, is situated right in column of the object, if the solution of the former does not proceed. Reference: Bodart and Vanderdonck [1993] p. 128 rec 7	Graduated

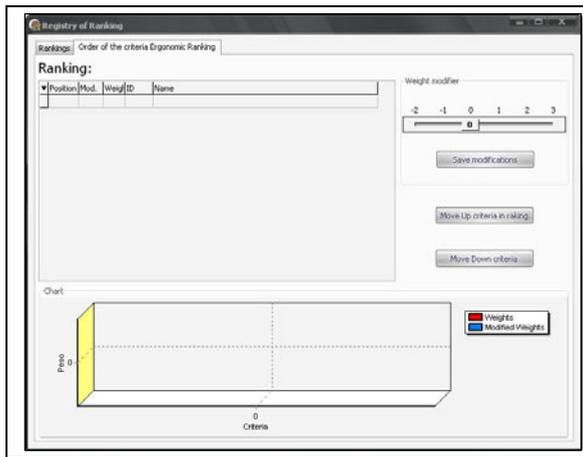
At the bottom left of the table, there is a 'Create' link.

Fig. 2. Question Presentation Window

Also, Figure 3 presents a window that presents the ranking of the Ergonomic Criteria according to a definition proposed by the evaluators and also according to the ranking of the answers presented by the ErgoCoIn's users.

We believe that using this new version we can get more efficacy and efficiency on web sites' usability evaluation. Therefore, the reports generated may present meaningful data to designers, what include usability taxes and metrics that may help designing new products or improving the ones that are already being used and have being evaluated.

We have performed some validation tests of this new version, comparing it with the application of "traditional" usability methods and also with the version 1.0. The results obtained indicate that this new approach is more attractive as it presents more conclusive and meaningful reports.



**Fig. 3.** Screen Aimed at Receiving Definitions Concerning Relative Ergonomic Criteria Importance

Now, to validate this new version of ErgoCoIn we must:

1. define a group of web sites to be evaluated. They must belong to different application domains and need to have specific well defined tasks to be evaluated (ie, tasks that have initial and final web pages well defined);
2. define a group of evaluators, aiming to have different skills and capabilities when interacting with the web sites under evaluation;
3. prepare and perform several usability evaluations in order to collect data to calculate usability taxes and metrics;
4. perform traditional usability evaluations of these web sites; and
5. compare the results obtained by both using the ErgoCoIn version 2.0 and the traditional usability evaluation methodologies.

These steps are proposed as future works to improve and validate the ErgoCoIn version 2.0.

## References

1. Downton, A.: *Engineering the Human-Computer Interface*. McGraw-Hill, London (1992)
2. ISO/DIS 9241. *Dialogue Principles in Guiding the Evaluation of User Interfaces, Guidance on Usability* (1997)
3. Brajnik, G.: *Automatic Web Usability Evaluation: What Needs to be Done?* In: 6th Conference on Human Factors and the Web, Austin, Texas, USA (2000)
4. Ivory, M.Y., Hearsstam, M.A.: *The State of the Art in Automating Usability Evaluation of User Interfaces*. *ACM Computing Surveys* 33(4) (December 2001)
5. Nielsen, J.: *Designing Web Usability*, Editora Campus, p. 416 (2000)
6. Storrs, G.: *The Notion of Task un Human-Computer Interaction*. In: HCI 1995- 10th Annual conference of the British Human-Computer interaction Group. University of Huddersfield, UK (1995)
7. Winckler, M., Palanque, P., Farenc, C., Pimenta, M.: *Task-Based Assessment of Web Navigation Design*. In: Proc. of TAMODIA 2002, Task Models and Diagrams for User Interface Design, Bucharest, Romania (2002)
8. Scapin, D.L., Leulier, C., Vanderbonckt, J., Mariage, C., Bastien, C., Palanque, P., Farenc, C., Bastilde, R.: *Towards Automated Testing of Web Usability Guidelines*. In: *Tools for Working with Guidelines*, pp. 293–304. Springer, London (2001)
9. Scapin, D.L., Bastien, J.M.C.: *Ergonomic Criteria for Evaluating the Ergonomic Quality of Interactive Systems*. *Behaviour and Information Technology* 16(4/5) (1997)
10. Molich, R., Bevan, N., Curson, I., Butler, S., Kindlund, E., Miller, D., Kirakowski, J.: *Comparative Evaluation of Usability Tests*. In: *Proceedings of the Usability Professional's Association Conference* (1998)
11. Wammi: *Website Analysis and MeasureMent Inventory (Web Usability Questionnaire)*, (n.d.) (2005), <http://www.ucc.ie/hfrg/questionnaires/wammi> (accessed in 2009)
12. Cybis, W.A.: *ErgoManager: a UIMS for monitoring and revising user interfaces for Web sites*. Rocquencourt: Institut National de Recherche en Informatique et en Automatique (2005) (Research report)
13. Keynote, A. *Research Manager Clickstream*, [http://www.keynote.com/solutions/solutions\\_cem\\_research\\_mgr\\_clickstream.html](http://www.keynote.com/solutions/solutions_cem_research_mgr_clickstream.html) (accessed in 2010)
14. Sommerville, I.: *Software Engineering*. Prentice-Hall, Englewood Cliffs (2009)
15. Morandini, M., Cybis, W.A.: *Definition and Prototyping of ErgoMonitor*. In: *Proceedings of International Conference on Human-Computer Interaction, Creta*, vol. II, p. 1396–1399 (2003)
16. Cybis, W.A.: *UseMonitor: suivre l'évolution de l'utilisabilité des sites web à partir de l'analyse des fichiers de journalisation*. In: *18eme Conférence Francophone sur l'Interaction Humain-Machine, Montréal. Actes de la 18eme Conférence Francophone sur l'Interaction Humain-Machine*, vol. 1, pp. 295–296. ACM - The Association for Computing Machinery, New York (2006)