

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/221565488>

# Website Evaluation: A Usability-Based Perspective

Conference Paper in Lecture Notes in Computer Science · November 2001

DOI: 10.1007/3-540-38076-0\_15 · Source: DBLP

CITATIONS

36

READS

2,391

4 authors:



**Nikolaos M. Avouris**  
University of Patras

288 PUBLICATIONS 3,900 CITATIONS

[SEE PROFILE](#)



**Nikolaos Tselios**  
University of Patras

115 PUBLICATIONS 1,384 CITATIONS

[SEE PROFILE](#)



**Christos A. Fidas**  
University of Patras

92 PUBLICATIONS 1,041 CITATIONS

[SEE PROFILE](#)



**Eleftherios Papachristos**  
University of Patras

15 PUBLICATIONS 122 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Facebook use, personality characteristics and academic performance [View project](#)



Web 2.0 in Education [View project](#)

# Website Evaluation: A Usability-Based Perspective

Nikolaos Avouris, Nikolaos Tselios, Christos Fidas, and Eleftherios Papachristos

Human-Computer Interaction Group  
Electrical & Computer Engineering Department, University of Patras  
GR-265 00 Rio-Patras, Greece  
{N.Avouris,nitse,fidas,papachristos}@ee.upatras.gr

**Abstract.** The usability is recognized as an important quality factor of any modern website. In this paper, techniques for usability evaluation of a website are proposed and their use is described. The results of application of these techniques are discussed in the frame of the design and evaluation of a web portal, serving an Academic Department. The applicability of the developed techniques in a wide area of web-based applications and their importance in the context of today's web applications development is also discussed in this paper.

## 1 Introduction

The evaluation of a website is an important phase of the development cycle, often overlooked during the fast-paced development processes adopted by modern web applications developers. This is more important today when the web becomes gradually a platform for deployment of complex applications of increased interactivity, a front end for databases and corporate information systems. This new use of the medium increases the importance of usability, as the web is used for accomplishment of complex tasks, like learning, retrieving information, interacting and collaborating with peers [1]. The effective accomplishment of such tasks has to be proved during the design of web applications. According to Ivory and Hearst [2], many web sites today are characterized by low usability, while millions of new web sites are expected to become operational during the next years, further intensifying the problem.

Today's highly interactive web applications tend to adopt interaction styles borrowed from traditional software. This is not however always acceptable, since the web poses special requirements that need to be taken in consideration [3]. For instance, the web users' characteristics are often not well known beforehand or can vary considerably. Also the computing platforms, network bandwidth and access devices used can be very different. For example, users can access a certain application through wap-enabled devices using 9,6 kbps connection, through typical modems of 33,6-56 kbps or through high bandwidth connections allowing few hundred kbps or higher. These can affect considerably the user interaction characteristics.

According to Nielsen [4], the highly quoted user-centred design methodology is considered applicable in this new context. The principal characteristics of this ap-

proach, as described by Theng et.al [5], are: (i) Interactive systems design should be based on a systematic user requirements capturing and recording of possible ways in which the users interact with the application. (ii) Design of highly interactive applications needs to be an iterative process. This approach, if adopted, brings iterative usability evaluation [6,7] at the center of the design process. However, involvement of users and setting up of usability testing experiments, advocated by user-centered design, is often considered a burden to the design team. Despite this, the need for such design approaches is widely recognized by the research community. Additionally there is an increased need to adapt the established usability testing techniques [6] (application inspection methods [4,8], observation methods [9], questionnaires and interviews [10]) to the requirements of web usability evaluation and to accelerate the often tedious process they propose [11,12].

In this paper, we report our experience with a website usability testing. The case study discussed relates to a web portal that has been designed and is in operation during the last years, to serve the academic community of the Electrical and Computer Engineering Department of the University of Patras in Greece ([www.ee.upatras.gr](http://www.ee.upatras.gr)). This application, typical of many similar University Departments, offers information on the activities of the Department to the public and the members of the Department and includes some more interactive components like support to the students for building their personal semester timetable and exams schedule.

A number of usability tests of this application have been conducted recently. These were based on adapted techniques proposed in the literature: (a) an experiment involved on-line questionnaires filled by the users, (b) heuristic evaluation by usability experts in the lab, and (c) user observation techniques and analysis of usage log files. The conducted experiments are presented and discussed. The findings are compared and their usefulness in the design of web applications is discussed. The re-design of the portal following these evaluation studies is also briefly described. Wider implications of the multi-faceted usability evaluation framework described in design of web applications are discussed in the last part of the paper.

## 2 Case Study: Usability Evaluation of an Academic Website

The web site of the ECE Department of the University of Patras ([www.ee.upatras.gr](http://www.ee.upatras.gr)) is a bi-lingual portal (in English and Greek) that has been developed with the objective to promote the Department and to provide information about its research and educational activities while at the same time to serve the Department students and faculty members. This latter use necessitates some interactive components, which need special attention during usability evaluation. The eight areas covered by the portal relate to:

- (a) Information on the Department history and structure,
- (b) Academic matters (courses offered, curriculum content etc.),
- (c) Research activity (laboratories, publications, results),
- (d) Personnel (faculty and researchers information, personal web pages etc.),
- (e) Additional information (the area, access instructions etc.),

- (f) Internal pages (password controlled services) and
- (g) Web site information (contact info, sitemap etc.).

One of the provided services, which has been the focus of evaluation, as discussed in the following, is the *search engine* of the site. The user of the search facility can select one of four information areas (faculty, graduate students, personnel, curriculum subjects) and use keywords for searching information in each of these areas. Another service of the site that has been evaluated is the *personal timetable* and *exam schedule*. Users can select, out of a menu, the courses or the exams to which they have been registered for and request their personal course timetable or exams schedule to be shown on the screen. In figure 1 typical interaction screens of the personal exam schedule application are shown. The courses/exams can be selected from an alphabetical list, attached to a combo-box, as shown in Fig. 1.

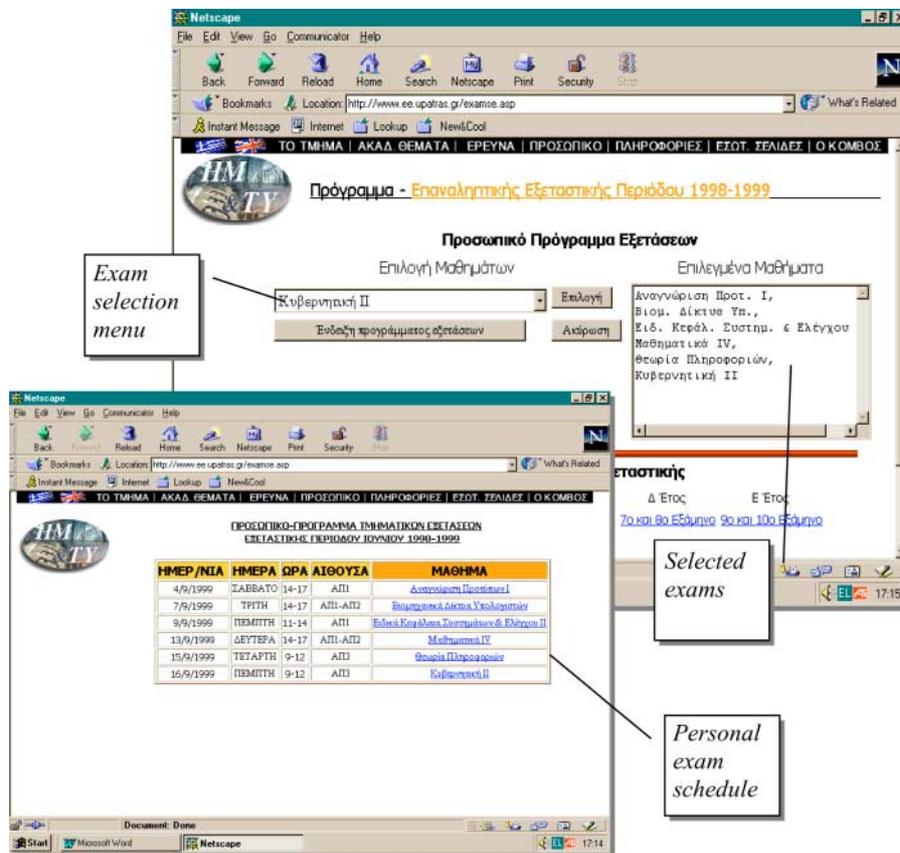


Fig.1. The personal exam schedule service (September 1999 exams shown)

### 3 Usability Evaluation Experiment

A series of usability evaluation experiments took place, after the web site had been in operation for two years. The experiments are presented in this section. In the final section of the paper a discussion on the findings is included.

#### 3.1 Questionnaire-Based Evaluation Studies

Questionnaires have been extensively applied in usability evaluation experiments [10, 11]. In our case two separate experiments have been conducted using this technique.

*On-line questionnaire.* The first one involved an on-line questionnaire, particularly useful in web usability evaluation when the users are distant, an extract of which is shown in figure 2. The questionnaire contained a number of simple questions requesting the view of the users on the quality of the provided services and the utility of the web site. The *zoomerang* ([www.zoomerang.com](http://www.zoomerang.com)) service was used for compiling the questionnaire and analyzing the responses.

The questionnaire was made of four parts. The first part requested the view of the users in relation to the *quality of the application* (ease of navigation, aesthetic design, quality of content and usefulness of services). The second part requested *frequency of use* and the view of the users in relation to specific site parts and services (personal course timetable, curriculum material, faculty/personnel, department info, diploma thesis information, news service). The requested answers in these two parts of the questionnaire were expressed in a multi-point scale ranging from 1 (inadequate) to 5 (excellent). Similar scale was used for the frequency of use questions. The third part contained questions on the *context of use*, like type of browser most often used, screen resolution, bandwidth etc. No personal data were requested. Finally, the fourth part of the questionnaire contained open questions on the three more *serious problems* encountered and the three more *useful services*. No specific suggestions were made to the users in this part of the questionnaire.

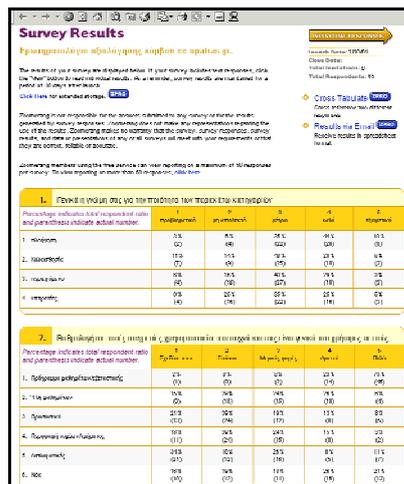


Fig. 2. Extract of the on-line questionnaire

**Table 1.** Evaluation of quality of design through online questionnaire

<i>Site feature</i>	answers					Mean value	Stand. Dev.	Mean error	Confidence limits
	Inadequate	poor	average	good	excellent				
<i>Navigation</i>	3%	6%	35%	44%	10%	<b>3.51 (average/good)</b>	0.88	0.1121	3.292-3.740
<i>Aesthetic design</i>	11%	14%	40%	29%	5%	<b>2.91 (average)</b>	0.94	0.1201	2.679-3.160
<i>Content</i>	6%	16%	43%	29%	3%	<b>3.06 (average)</b>	0.92	0.1189	2.828-3.303
<i>Services</i>	6%	25%	35%	25%	5%	<b>2.96 (average)</b>	0.99	0.1280	2.711-3.223

For twenty (20) days the questionnaire was linked through a visible button to the home page of the site. During this period sixty five (65) users volunteered a response. An overview of the responses to the first two parts of the questionnaire is shown in Tables 1 and 2. According to these tables, navigation was considered good, while aesthetic design, in-formation content and the quality of the provided services received an average mark.

The Usefulness/frequency of use questionnaire demonstrated the high value of the personal timetable and exam schedule service, while the other services received below average marks, probably not considered useful and adequately designed. The third part of the questionnaire revealed useful information about the users of the site, like browsers used and connection bandwidth. For instance from the answers of this part it was deduced that more than 60% of the users connect to the site through modem, thus special attention should be paid not to include heavy web pages in the site, while all the responders used computing equipment with screen resolution equal or higher to 800X600 pixels, suggesting that the web pages design should be adequately adjusted to make for more efficient use of screen space.

Finally the replies to the open questions of the fourth part of the questionnaire, shown in Table 3, identified the need for new services like links to library, course marks on line, student magazine etc, while revealed some usability problems, like the low visibility of the link to web servers of laboratories, a service that despite the fact that it was already provided, it was proposed as a new service in Table 3.

**Table 2.** Frequency of use according to online questionnaire

<i>Site part</i>	<i>Frequency of use</i>					Mean value	Stand. Dev.	Mean error	Confidence limits
	Never	Rarely	Sometimes	Frequently	Very Frequently				
<i>Personal timetable/ exams schedule</i>	2%	0%	3%	23%	73%	<b>4.64 (frequently/very frequently)</b>	0.70	0.08	4.466-4.824
<i>Curriculum material</i>	15%	29%	24%	26%	6%	<b>2.80 (sometimes)</b>	1.17	0.14	2.509-3.104
<i>Faculty</i>	21%	39%	19%	13%	8%	<b>2.48 (sometimes)</b>	1.19	0.15	2.180-2.780
<i>Department</i>	18%	39%	24%	15%	3%	<b>2.45 (sometimes)</b>	1.05	0.13	2.188-2.730

**Table 3.** Proposed new services and identified problems according to the online questionnaire

Proposed new services		Most severe problems	
<i>On line Grades for individual students</i>	35,7%	<i>Late update</i>	42,9%
<i>Richer news services including student news</i>	32,1%	<i>Out of date layout</i>	23,8%
<i>Electronic delivery of applications to secretary</i>	28,6%	<i>Lack of detailed news services</i>	19,0%
<i>Relevant links to library</i>	21,4%	<i>Small number of services delivered</i>	9,5%
<i>Links to lab/department web servers</i>	17,9%	<i>Lack of links to lab/department web servers</i>	9,5%
<i>Relevant course material download</i>	14,3%	<i>Greek encoding problems in some pages</i>	4,8%
<i>Detailed exams material for courses</i>	14,3%	<i>Slow response</i>	4,8%
<i>On line student magazine</i>	10,7%	<i>Problems in search</i>	4,8%

**Table 4.** Replies of WAMMI questionnaire in a strongly disagree (1) to strongly agree (5) scale

<i>This web site has much that is of interest to me:</i> 3.83 (agree)
<i>It is difficult to move around this web site:</i> 2.33 (disagree)
<i>I can quickly find what I want on this web site:</i> 3 (neutral)
<i>This web site seems logical to me:</i> 4 (agree)
<i>This web site needs more introductory explanations:</i> 3.16 (neutral)
<i>The pages on this web site are very attractive:</i> 2.83 (neutral)
<i>I feel in control when I'm using this web site:</i> 3 (neutral)
<i>This web site is too slow:</i> 2.16 (disagree)
<i>This web site helps me find what I am looking for:</i> 3.33 (neutral)
<i>Learning to find my way around this web site is a problem:</i> 2.16 (disagree)
<i>I don't like using this web site:</i> 2.66 (neutral)
<i>I can easily contact the people I want to on this web site:</i> 4.33 (agree)
<i>I feel efficient when I'm using this web site:</i> 3 (neutral)
<i>It is difficult to tell if this web site has what I want:</i> 2.83 (neutral)
<i>Using this web site for the first time is easy:</i> 3.33 (neutral)
<i>This web site has some annoying features:</i> 3 (neutral)
<i>Remembering where I am on this web site is difficult:</i> 1.66 (disagree)
<i>Using this web site is a waste of time:</i> 1.83 (disagree)
<i>I get what I expect when I click on things on this web site:</i> 2.83 (neutral)
<i>Everything on this web site is easy to understand:</i> 3.16 (neutral)

*The WAMMI questionnaire.* A second experiment involving a questionnaire was directed to users who participated in the users observation experiment described in section 3.3. In this case the users (six experienced users of the system) were asked to fill a Wammi inspired questionnaire [13]. The requested replies were in Likert-type scale (strongly agree / agree / neutral / disagree / strongly disagree), though which the users expressed their attitude towards the statements. This kind of questionnaire, filled by users during a usability experiment can be used for checking their attitude and can be complementary to the usability evaluation experiment. The average values of the received responses are shown in Table 4. Comparing this questionnaire with the online one presented above one can verify that the results are not contradictory.

*Validation of questionnaires with logfiles.* Finally, a cross checking of the findings of the questionnaires with log files collected by the web server was performed. This identified the frequency of use of parts of the site as well as characteristics of accessing equipment (e.g. browser version). Data from a period of two months has been used. The study of the log files confirmed the validity of the filled questionnaires. For instance the browser type data of the log file and the questionnaire were almost identi-

cal. However the emphasis is different, the log files do not present the views of the users, while provide some additional information like patterns of use during the day/week/year, referring domains etc, which can be useful complementary material for user profiling ([14],[15]). Development of automatic techniques for analysis of these log files and use of them in a complementary way in the evaluation is a matter of further research.

### 3.2 Heuristic Evaluation

The second evaluation technique involved eleven (11) evaluators, knowledgeable in user-centered design, with high web usage experience. They applied an inspection technique, called *heuristic evaluation* [4] using heuristic rules that govern good interface design. This is a well-known, highly appraised experts-based evaluation technique, which has been successfully used in many similar experiments [16,8].

*Method of the study.* During the first phase of the study, the evaluators were encouraged to navigate the application and carefully validate the implementation of each heuristic rule. Evaluators were asked to simulate execution of representative tasks during this phase. When a rule violation was detected, the evaluator identified where the violation occurred, and proposed likely solutions. Each evaluator filled a report describing his/her findings. Evaluators were also provided with an evaluation sheet, in which they could quantify their subjective judgment, by assigning marks expressed in a 5 mark scale, indicating degree of conformance with each particular rule. The range of assigned values was 1 (inadequate) to 5(excellent). A relevance weight ( $w_i$ ) was also assigned to each rule, according to the expert's opinion. Quantitative results of the evaluation are presented in table 5. Factor-weighted score,  $u$ , derives according to

$$u = \frac{\sum_{i=1}^{10} r_i \cdot w_i}{\sum_{i=1}^{10} w_i} \quad (1)$$

where  $r_i$  is the rule rating and  $w_i$  is weight for each heuristic rule  $i$ . The overall usability score was average to good (3.38 excluding rule weight factors, and 3.42 including weight factors). 7 out of 10 heuristic rules ranked with a relative high grade. Rules h7, h9 and h10 received average grade, indicating usability problems related to the specific rules as discussed in the following.

*Heuristic Rule h7. (Flexibility and efficiency of use)* The system in general does not provide shortcuts to facilitate user's tasks especially in complex processes such as Searching and Personal Timetable. For example, frequently searched terms are not provided in a form of ready-to-select links. Also users cannot make multiple-course selections in Timetable building, such as all the courses of a specific semester. Additionally to this no specific reference to post-graduate courses was allowed.

**Table 5.** Average subjective score, weight factor, and total errors found for each heuristic rule

Heuristic Rule : <sup>1</sup>	h1	h2	h3	h4	h5	h6	h7	h8	h9	h10
Score (mean value)	3.63	4.45	3.63	4.18	3.27	3.54	2.55	3.72	2.63	2.19
Stand. deviation	0.36	0.28	0.24	0.37	0.38	0.36	0.31	0.19	0.27	0.22
Relevance (mean value)	4.54	4.45	3.45	4.45	4.27	4.18	3.18	3.81	3.81	3.72
Stand. deviation	0.28	0.24	0.34	0.24	0.27	0.35	0.35	0.35	0.40	0.33
Number of detected Errors	4	3	2	3	7	2	3	2	5	2

*Heuristic Rule h9. (Support the users to recognize, diagnose, and recover from errors)* The most severe problem is that the typical 404 page has not been replaced with a page showing the site map of the site, explaining that a system error occurred (possibly because of wrong URL), and providing a search link. Also some error messages do not constructively suggest solutions or indicate why the problem occurred.

*Heuristic Rule h10. (Help and documentation)* Even though the system can be used without a lot of online documentation as resulted from the user's observation experiment, at least some help should be provided in high interactivity tasks (search engine and Personal timetable). The help should be easy to search, focused on the user's task and explicitly list the steps to be carried out.

Overall thirty-three (33) usability flaws were unveiled during this study. Many of the usability problems were reported by more than one evaluators, confirming Nielsen's and Landauer's finding that four to five evaluators typically unveil about 80% of the overall usability problems [17].

### 3.3 User Observation

The third usability evaluation technique used involved observation of typical users, as they interacted with the application, executing typical tasks. The experiment took place in a usability laboratory. The users were six students of the Department, 4 undergraduate and 2 post-graduate, 5 male, 1 female, aged 23-28. They were all experienced users of the Internet, which they use daily in their activities. The equipment used in the study was a Personal Computer (Pentium III, 600 MHz, 128 Ram, 17" monitor, 1024\*768 screen resolution) with good connection to the web site through the high band University network.

At the beginning of the experiment, the users were informed that they were not them-selves the subject of evaluation. Five tasks were carefully described to them (table 6), however, while the objective of each task was made clear, no precise instructions on its execution were provided in order not to condition the users' strate-

<sup>1</sup> The applied heuristic rules were the following: 1. Visibility of system status, 2. Match between system and the real world, 3. User control and freedom, 4. Consistency and standards, 5. Error prevention, 6. Recognition rather than recall, 7. Flexibility and efficiency of use, 8. Aesthetic and minimalist design, 9. Help users recognize, diagnose, and recover from errors, 10. Help and documentation

gies [9]. The users were provided with sufficient time to complete their tasks. The observers did not provide any additional information or support. Timing of tasks and observations on strategies and patterns of use were made by the evaluation facilitators.

First remark was that Tasks T2, T3 and T5 were completed promptly with no special difficulties by any of the students. This proved that the application was sufficiently well structured to support the users identify specific information items. In contrary tasks T1 and T4, proved to be time-consuming and cognitively demanding.

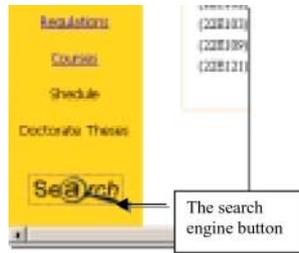
The developed strategies varied; often multiple browser windows were used in order to accomplish the tasks. Usability problems in the application design clearly emerged. In the T1 the courses were sorted alphabetically in the Personal Timetable page, no provision was made for alternative presentation (e.g. cluster per academic year). Also lack of redo and undo option hindered considerably task execution. A search option was not available in the same page. Availability of a single button for transfer of all courses of an academic year in the timetable would have considerably accelerated the process.

The search engine of the site was hidden behind a relevant icon, as shown in Fig. 3. This icon was not easily identified by the users as a button, since there was no familiar search textbox associated. So many users had difficulties in locating the Search engine of the site.

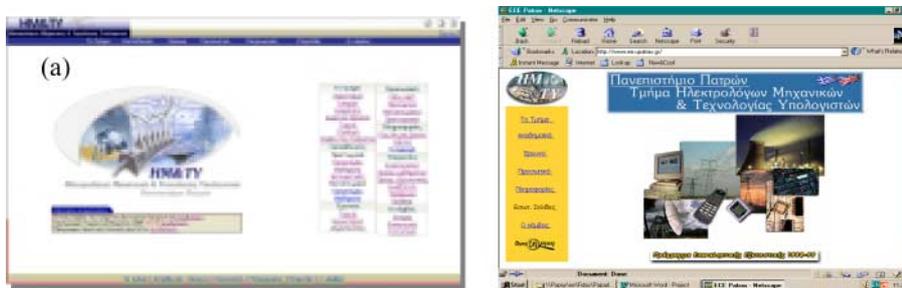
In order to accomplish task T4 the users needed to use the site Search engine. Some misunderstandings related with the use of this facility were identified during the study. The Search engine requires selection of a search category first, before inserting a keyword. There was no way to make a global search in all categories (course subjects, faculty, undergraduate students, post-graduate students etc.). Also the logic applied by the search engine was not made clear and was not explained. (e.g. Computer Networks searches for Computer OR Networks or the string "Computer Networks"). Finally, a subsequent search was not clear if it was made in the first search results or in the original search space. Some insecurity was observed to some users in relation to the validity of the search results.

**Table 6.** Description of tasks and time required to complete them

Task description	Average Time (sec)	Stand. Dev	Mean error
T1: To build a personal timetable for a student of the third year who is also registered to a limited number of courses of the 2nd year	266.33	54.58	22.286
T2: For a specific course to search for the course description and relevant handouts and other material	34.83	18.17	7.418
T3: To search for details of the CV of a faculty member (University and year of PhD award)	20.66	6.77	2.765
T4: To find all courses offered in the curriculum that involve computer networks	400.33	158.85	64.851
T5: To find information relating with access by public transport to the Department and the University.	19.16	7.521	3.07
<b>Total</b>	<b>741.33</b>	<b>196.71</b>	<b>80.305</b>



**Fig. 3.** The search button, through which the search engine of the site is invoked



**Fig.4.** The home page of : (a) the re-designed portal and (b) the original portal

One user did not manage to complete Task T4 for this reason. In summary 29 out of 30 tasks were successfully accomplished, however the difficulties encountered in tasks T1 and T4, as discussed here and the unexpected long times required for accomplishing the tasks, taking into consideration that expert users were involved, unveiled many limitations of the design.

#### 4 Website Re-design

The Academic Department web portal has been recently re-designed, taking in consideration the results of the usability evaluation studies described in the previous section. The new design is characterized by improved aesthetic appearance of the site, lower download times due to reduced multimedia content, following the user access data, more compact layout and more information content per webpage. At the same time the original structure and the content of the website were maintained.

A considerable modification was made to the *search facility* and the *personal timetable* and *exams schedule* services, following the remarks of the users and the observed difficulties in particular during the user observation experiments. For instance the selection screen for courses in the personal timetable service, in the new portal is done through a list of all available courses, grouped by year, as shown in figure 5, instead of an alphabetic list through a combo-box in the original design. This screen provides an overview of all courses. The user is provided with the possibility to select or deselect the courses from this view and then proceed with the timetable. This new design tackles the observed problems in the course selection task, discussed in section 3.



Fig. 5. The course selection screen in the personal timetable service

Table 7. Heuristic evaluation of the updated portal design

Heuristic rule	h1	h2	h3	h4	h5	h6	h7	h8	h9	h10
Score $r_i$ (mean value)	4,22	4,33	3,83	4,50	3,28	4,11	3,12	3,94	2,61	3,89
Relevance $w_i$ (mean value)	4,22	4,56	3,39	4,06	3,78	4,11	3,06	3,94	4,06	3,56

#### 4.1 Heuristic Evaluation of the New Portal Design

The new portal design has been recently evaluated, using the adapted heuristic evaluation approach discussed in section 3.2, in order to compare the findings of the evaluation with the evaluation study of the original design, discussed in section 3. The evaluators in this experiment were seventeen (17), all experienced in using the

evaluation technique and powerful web users. This was a different set of evaluators than those involved in the studies discussed in section 3.

The heuristic rules applied were those presented in section 3.2, while special attention was put in observing the evaluation protocol followed in the original evaluation study in order to be able to compare the findings of the two evaluation studies. A number of useful observations and suggestions were made by the evaluators. In table 7 the overview of the evaluation results is shown.

From the scores in table 7 and those of table 5, it was deduced that the mean value of the overall usability score of the updated design was 3.78, compared to 3.38 of the overall usability of the original design. This difference was considered significant, according to a pair wise t test ( $P(t)=0.036$ ).

A test was performed in order to compare the evaluators' attitude towards the heuristic rules used in these evaluation studies. So a pair wise t test was performed between the values of  $w_i$  (relevance of rule) for the two sets of evaluators of the specific website.

The mean value of the rules relevance in the two studies was proven not statistically significant ( $P(t)=0.1649$ ), an indication that the two independent groups of heuristic evaluators had similar attitudes towards the heuristic rules and their relevance to the specific site.

The difference between the two evaluation studies is also shown in figure 6. According to this figure, the score in heuristic rules 7 (flexibility) and 10 (help and support) has significantly improved, while that of rule 9 (error recovery) had similar value in both studies.

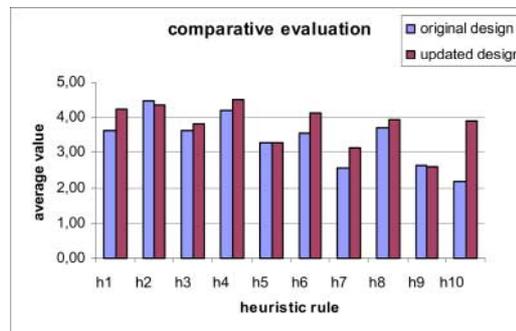


Fig. 6. Comparison of heuristic usability evaluation scores in the two portal designs

## 5 Discussion – Conclusions

Three complementary usability evaluation techniques were presented and their application in the usability evaluation of a web portal was described in this paper. All three presented approaches provided the design team with useful insight into the application and revealed many usability problems not previously detected.

The *on-line* and *Wammi* questionnaires helped us collect information relating to the users view on the supplied services and content of the site. The questionnaires are

techniques easy to implement, however the findings are mostly useful for identifying problem areas without suggesting solutions. Also crosschecking of the findings should be performed. The usage log files of the website can be valuable complementary information. Additionally, control questions incorporated in the questionnaire can be used for quality control of the supplied answers.

In addition, it should be observed that the questionnaires sometimes are not filled by representative part of the user population and can provide a distorted view of use. For instance, in our case the questionnaire was mostly filled by students and not faculty members, something reflected in the provided suggestions.

The second technique used was heuristic evaluation by experts. This technique requires more effort, since experts have to be engaged in experimentation. However the number of problems identified by this experiment, as presented in section 3.2, was impressive. A limitation was however that the identified problems were not deduced from realistic interaction patterns, but from dialogue flow analysis. This was due to the fact that this technique does not involve real users, but experts making hypotheses on user attitude.

The third experiment, involving real users closely observed as they executed typical tasks, revealed additional problems often overlooked by the designers and not identified by the previous two experiments. Typical examples of analysis of the observations made during this experiment were included in section 3.3. However this technique requires additional effort, as typical users have to be identified and actively engaged in the experiment, not an easy task in commercial and general-purpose applications. Also the experiment has to be made in adequately equipped environment and has to be well designed. Finally the effort required for analysis of the findings and their interrelation to design decisions is not trivial. So it is questionable if this technique can be fully implemented as part of the typical fast web application development process.

The findings of all three evaluation studies were transformed in new design requirements and were fed in the development team who recently updated the portal. This is the case with most web sites of this nature who need to go through a major redesign phase every two to three years. The new design was recently evaluated, by a new group of evaluators, using heuristic evaluation in the same way as in the evaluation study of the original design. The number of observations was in this case as high as in the original evaluation. By comparing the usability score of the two studies, the improvement in usability was considered statistically significant.

In conclusion, this study has shown us that by incorporating systematic usability evaluation studies in the development of the web portal, a significant improvement of the web site usability was achieved. We feel that an approach similar to the one described in this paper is indispensable part of a user-centered design and it should be followed in a rigorous way in websites where the customer's needs and desires need to be taken in consideration in great extend, like for instance portals of commercial nature.

In addition, the reported study showed that alternative usability evaluation techniques put different emphasis and provide the designer with different findings. So preferably a combination of them should be used in web application development. Given the importance of usability evaluation for development of this kind of applica-

tions and the effort required for most of usability analysis tests, it is worth investigating further techniques and supporting tools that automate parts of the described usability evaluation process.

## Acknowledgements

The web site described in this paper has been developed by a group of volunteers, members of the ECE Department. Financial support by the EPEAEK/ECE Department Curriculum Development Project and the PENED 99ED234 Project is acknowledged. Also the free service of Zoomerang of Market Tools Inc., used in the frame of the reported study, is also acknowledged.

## References

- [1] Shum S. B.: "The Missing Link: Hypermedia Usability Research & the Web", *ACM SIG-CHI Bulletin*, Vol.28, No.4, pp.68-75, 1996.
- [2] Ivory M.Y., Sinha R. and Hearst. M.A.: "Improving Web Site Design", *IEEE Internet Computing*, pp.56-63, March-April 2002.
- [3] Bevan N.: "Usability Issues in Web Site Design", *Proceedings UPA'98 Conference*, Washington DC, 1998.
- [4] Nielsen J.: "*Usability Engineering*", Academic Press, London, 1993.
- [5] Theng Y.-L., Jones M. and Thimbleby H.: "Lost in Hyperspace, Psychological Problem or Bad Design", *Proceedings 1<sup>st</sup> AsiaPacific Conference on HCI*, pp.387-396, 1996.
- [6] Avouris N.M.: "Introduction to Software Usability", *Proceedings 8<sup>th</sup> Panhellenic Conference on Informatics (Workshop on Software Usability)*, Vol.II, pp.514-522, Nicosia, November 2001.
- [7] Fitzpatrick R. and Higgins C.: "Usability Software and its Attributes: a Synthesis of Software Quality", *European Community Law and HCI, People and Computers XIII*, pp.3-22, Springer, London, 1998.
- [8] Levi D.M. and Conrad F.G.: "Usability Testing of World Wide Web Sites", *Research Papers of Bureau of Labor Statistics*, 1998. Available at <http://www.bls.gov>
- [9] Rubin J.: "*Handbook of Usability Testing*", Wiley, New York, 1994.
- [10] Feinberg S. and Johnson P.Y.: "Designing and Developing Surveys on WWW Sites Getting Feedback on your Web Site", *Proceedings 16<sup>th</sup> ACM Conference on Systems Documentation*, pp.38-42, 1998.
- [11] Kantner L. and Rosenbaum S.: "Usability Studies of WWW Sites: Heuristic Evaluation vs. Laboratory Testing", *Proceedings 15<sup>th</sup> ACM Conference on Systems Documentation*, pp. 153-160, 1997.
- [12] Keevil B.: "Measuring the Usability Index of your Web Site", *Proceedings ACM Conference on Human Factors in Computing Systems (CHI)*, pp.271-277, 1998.

- [13] Management N. and Kirakowski J.: “Web Usability Questionnaire”, 2000. Available at <http://www.nomos.se/wammi>
- [14] Hochheiser H. and Shneiderman. B.: “Understanding Patterns of User Visits to Web Sites: Interactive Starfield Visualizations of WWW Log Data”, *Proceedings ASIS'99 Conference*, 1999.
- [15] Pirolli P. and Pitkow J.E.: “Distributions of Surfers' Paths Through the World Wide Web: Empirical Characterization”, *World Wide Web*, Vol.2, No.1-2, pp.29-45, 1999.
- [16] Avouris N.M., Tselios N.K. and Tatakis E.C.: “Development and Evaluation of a Computer-based Laboratory Teaching Tool”, *J Comp. Applications in Eng. Education*, Vol.9, No.1, pp.8-19, 2001.
- [17] Nielsen J. and Landauer T.K.: “A Mathematical Model of the Finding of Usability Problems”, *Proceedings ACM/IFIP INTERCHI'93 Conference*, pp.206-213, Amsterdam, The Netherlands, April 1993.