

Experimental Evaluation of an Instructional Supporting Tool in Distance Learning

Athanasios Karoulis, Ioannis Stamelos and Lefteris Angelis

Dept. of Informatics, Aristotle University of Thessaloniki, Thessaloniki, Greece // Tel. 0030/2310/99.19.10 // Fax. 0030/2310/99.84.19 // athanasios@karoulis.gr // stamelos@csd.auth.gr // lef@csd.auth.gr

ABSTRACT

This paper describes a controlled experiment concerning the use of a learning aid during an open and distance learning (ODL) course. The core issue of investigation is whether this instructional aid can support, guide, and scaffold the distant student in his/her study. For this purpose, a controlled experiment was conducted with the participation of 191 undergraduate students at the department of informatics at a university in Greece. The considered domain was two lessons concerning human–computer interaction (HCI), the first in usability engineering and the second in interface evaluation methodologies. A test session was also conducted to collect data on the assessment of the effectiveness of the proposed tool. Descriptive statistics as well as a variety of statistical methods have been applied to the collected data in order to test the research hypotheses. The results have shown a statistically significant difference in performance for the student group that used the tool. In order to validate these results, a validity evaluation was also considered and presented. Finally, concerns about the application of the tool in a broader context and further research on the area are also presented.

Keywords

Lesson sheets, Open and distance learning, Controlled experiment

Introduction

Theoretical framework of the study

The constructivist model of learning assumes that knowledge is not transferred but is created by the learner, with the instructor as facilitator (Belanger & Jordan, 2000). Most recent approaches adopt the student-centered approach, which is clear and understandable (Lowman, 1981), is responsive to the ways in which students learn and communicate (Kolb, 1984), acknowledges students' interests and motivations (Forsyth & McMillan, 1991), and focuses on the explicit needs of learners for meaningful and timely feedback (van Houten, 1980).

All paradigms in this area suggest that ad-hoc prepared instructional material, either in open and distance learning (ODL) or in traditional approaches, has to follow specific guidelines in order to be characterized as sound (Georgiadou & Economides, 2000). Aims and targets, keywords, short chapters and paragraphs, simple language, explanation of difficult points, and exercises and activities are only some of the guidelines. In addition to these, new educational approaches and methodologies have evolved to integrate this material more smoothly into the instructional procedure. Furthermore, various learning aids have been proposed, such as study frameworks, time scheduling, or summaries, because learning issues are, after all, complex and highly case sensitive (Squires & Preece, 1996). So, there is a broad consensus that instructional tools aim to support the student in his/her work, to confront eventual learning difficulties, and to facilitate the students in personalizing the offered information.

Existing research and motivation of this study

In the described context, a new instructional tool, simply called the lesson sheet, has been proposed and thoroughly described in Karoulis, Demetriades, and Pombortsis (2004a). The investigation of the potency of this tool underwent three phases. The first one concerned a qualitative survey, which was applied in traditional classes and was published in the aforementioned study. The second evaluation was a controlled experiment concerning the application of the tool in both a traditional and a hybrid class, and is thoroughly described in Karoulis, Stamelos, Angelis, and Pombortsis (2004b). The work presented here concerns the third phase of the study, namely, to modify the proposed tool to make it suitable for ODL, to apply it in the framework of a pure distance-learning course, and to evaluate its potency and effectiveness in this context. So, a new controlled experiment has been organized and performed in order to elucidate this question.

In short, undergraduate students of the department of informatics studied, through distance learning, two lessons on usability engineering within the framework of a course on software engineering. The students were divided into two groups: a control group, consisting of students who didn't use the modified lesson sheet, and an experimental group, consisting of students who did use the modified lesson sheet. At the end of the predefined period, an in-person examination took place, with identical questions for both groups. The collected data was analyzed statistically. The results showed a statistically significant difference in the performance of the group of students who used the tool. As the main focus of this paper is to describe the pedagogical findings and to raise discussion based on these findings, this result points towards the application of a simple but valuable tool for supporting, guiding, and scaffolding the distant student during his work at home.

We emphasize at this point that prior pedagogical findings of the application of the tool in traditional and hybrid (blended learning) settings have already been published in Karoulis, Demetriades, & Pombortsis (2004) and Karoulis, Stamelos, Angelis, & Pombortsis (2004) so present work aims to add the findings of its application in ODL settings, as subsequently described. In order to be precise in this evaluation, we followed a rigorous methodological framework in this paper. While this approach is encountered often in other disciplines, it is greatly neglected in contemporary ODL research.

Lesson sheet

The proposed tool has its origins in the domain of open learning. In table format, it is divided into rows and columns. The left-hand column contains the outline of the lesson in paragraph form, while the right-hand column provides a variety of context-sensitive information that could be a summary of the subject under consideration, such as charts and graphs, pictures, or other material related to the discussed issues. In addition, adequate free space allows students to personalize the sheet with his/her own notes — information that should have meaning to students when they work at home.

The basic construction guidelines of the lesson sheet are listed below:

- The lesson sheet uses the shape of a table to frame and structure the offered information.
- The left-hand side contains, the information presented. This provides an outline for every lesson.
- There must be adequate recall of situated knowledge: an already known definition, image, graph, etc.
- On the right-hand side there is space for the student to note whatever he/she wants relevant to the corresponding lesson to thus personalize the lesson.
- Images, figures, charts, and any other appropriate means are included to help students visualize the information.
- Notes that have nothing to do with personalization, such as Internet links, tables, and additional or external information and resources, are provided on the sheet in order to accelerate the procedure and diminish the distraction that occurs during the students' note-taking time.
- The left-hand side includes the necessary self-assessment exercises and the activities that the student has to perform during the lesson, while the corresponding space on the right is usually reserved for the answer. Exercises gradually increase in difficulty.
- It must provide feedback on the self-assessment exercises so that the student can correct his/her mistakes, by various means, such as pinpointing the relevant sections on the instructional material or suggesting further readings or external sources.
- Lesson sheets contain text and theory of no more than 2 to 3 paragraphs in length for every section. In addition, an alternative option such as an exercise, an external resource, or a student activity is provided.
- The use of the sheet in practice has shown that 4 to 5 pages are sufficient for a 2-hour lesson. In addition, good image and print quality are important.

Appendix A is a simplified sample lesson sheet, which clarifies the basic structure. The sample is based on the lesson sheet used during this experiment.

Experiment

Since this study concerns a controlled experiment, emphasis was given to two areas. First, experimental settings were made as realistic as possible with respect to those in practice. To fulfill this requirement, the lesson sheets were

applied to real university-level, undergraduate lessons, delivered through distance learning. Furthermore, an external motivation existed because the test score counted towards each student's final score. In this way, the students perceived this series of sessions to be part of their syllabus. Second, planning, operation, and analysis were done according to the methods proposed in the literature, such as in Wohlin et al. (2000). Consequently, the procedure described adheres to these guidelines.

Definition

To define the experiment, the goal-question-metric model by Basili & Weiss (1984) and Basili & Rombach (1988) was used. The formal definition of this experiment is as follows:

Analyze	<i>the application of the lesson sheets</i>
For the purpose of	<i>assessment</i>
With respect to their	<i>effectiveness during the instructional procedure</i>
From the point of view of	<i>instructors and instructional designers</i>
In the context of	<i>distance learning university level lessons on usability engineering</i>

Planning

The context of the experiment is determined in this section. This includes personnel and the environment, the selection of variables, measurement scales, the statement of the hypotheses, and so on.

Context selection

To achieve the most general results, the experiment should be executed in real settings. However, due to various constraints, this is often not feasible, so a cheaper alternative is chosen, such as using a restricted version of the real environment or students as experiment subjects. Such approaches are cheaper and easier to control, yet they seldom address real problems and are more directed to a certain context. So, the context of the experiment must be defined according to following four dimensions (Wohlin et al., 2000):

- **Online vs. off-line.** Online, since it was performed during the semester's lessons.
- **Student vs. professional.** Students are the final user population to use the tool, so "students" and "professionals" in this context are equal.
- **Pretend vs. real problem.** The experiment addresses a real problem under real circumstances (real lessons).
- **Specific vs. general.** This investigation concerns a specific problem, namely, the subject of usability engineering. Generalization to other domains may be made (if at all) after having considered the external validity of the experiment. This issue will be addressed later.

Selection of variables

The independent variable (factor) is the use of the lesson sheet. It is of nominal type, with two possible values (yes & no).

The dependent variable is the students' performance on the taught domain, measured in terms of a grade on a test. This variable is measured in ratio scale, from 0 to 150, and is operationalized in terms of 15 questions depicted in the examination sheets of the final test.

Null hypothesis

H_0 : *There is no difference in the students' performance on the taught domain due to the use of the lesson sheets.*

The alternative hypothesis is consequently:

H_a : *There is a difference in the students' performance on the taught domain due to the use of the lesson sheets.*

Selection of subjects

Two independent groups were formed from 191 participating students attending the fifth semester at the department of Informatics at the Aristotle University of Thessaloniki, Greece (AUTH). The students were all volunteers, and the grade they would acquire at the scheduled final test would count towards a bonus of up to 10 percent above their final grade. They were divided into a control group of 96 students and a treatment group of 95 students. Students had been previously grouped based upon their choice of percentage in the following statement, which appeared in the header of examination document: “Percentage of the use of the tool during your study: 0%, 10% ... 90%, 100%.” According to their declaration, students who had used the sheet less than or equal to 50 percent were classified as the control group, whereas students who had used it more than 50 percent were classified as the treatment group. The final classification was balanced (although this was not intended): 96 vs. 95 students.

Educational framework

Initially, the supporting website was set up. Through this site, the following educational material, consisting of four main learning files and some supporting material, was delivered to the students:

- The lesson sheet, a file named “Guide.pdf”
- The two lessons files, “HCI01.pdf” and “HCI02.pdf”
- A third learning file, “Presentation.zip,” an Authorware file. As stated to the students, “this would be the file a traditional teacher would use to present the material, if these lessons were in traditional form.”
- A shareware book (*Task-Centered User Interface Design*, by Lewis & Rieman, 1994) and some relevant Internet resources (provided as additional, voluntary instructional material).

Experimental design

Lessons

The students participated in, over distance, the equivalent of two 3-hour lectures (one per week) concerning two selected aspects of usability engineering (UE): introduction/definitions of UE, and evaluation.

The studying material consisted of two pdf files, one for each “lecture.” There was also a multimedia presentation, presenting all aspects the text files in a multimedia way. External resources, such as additional text books and relevant websites were also provided. The students had also the option to contact a mentor (one of the authors) if they had questions regarding their study. In conclusion, the design aimed to imitate a self study in a distance environment, corresponding to two real lessons and providing adequate support to the student.

The main supporting and guiding file was the modified and enriched lesson sheet, also provided at the supporting website. The students had the choice of whether to use it or not. They only had to declare this percentage of usage at the final test, in order to be classified into a group, as already described.

Tool

The lesson sheet was a detailed guide that accompanied students through their study, pinpointing pages to read, providing self-assessment exercises (with answers at the end of the sheet), encouraging them to visit the supporting Internet resources (where relevant), providing explanations on unclear or difficult aspects of the material, and providing pictures or diagrams where appropriate. In other words, the sheet was designed to be scaffolding and a guiding factor for the student who would rely on it. Obviously, this last claim is the main one to be proved through this experiment.

Examination

The final test took 2 hours and consisted of an examination sheet with 15 multiple choice questions to facilitate the statistical analysis.

The experimental groups were, as already stated, randomized and almost balanced with 96 and 95 participants respectively. No blocking of any subject with respect to other factors was made.

The purpose of this design was to depict and measure the impact of the learning sheet on the students' performance. In this case, it was most appropriately recorded through a written test, as there were a large number of participants, which prohibited other approaches such as interviews or oral examinations. Note that there were no criteria such as success or failure for the students, only the acquired grade in a scale from 0 to 150, as there were 15 questions in the examination sheet, each worth 10 points.

Other materials and information

In addition to the lesson sheets described above, students were also supplied with the following:

- Two distance lectures, as Adobe pdf files, which concerned UE basics and evaluation methodologies. These lectures were chosen because they provide many new notions and include simple as well as complex material. Students in both groups received the same quality and quantity of information.
- Instructions to students on the website, the time scheduling etc. It was emphasized that the accompanying lesson sheet was a valuable, yet optional, aid in their study. However, as results showed, only half of the students used it in the way it was intended to be used.
- A presentation of the issues discussed in the pdf files. The Authorware multimedia-authoring tool was preferred here over PowerPoint because of its augmented multimedia capabilities.
- Additional studying material, such as a shareware book and some Internet resources, as already described.
- The final examination.

Guidelines to subjects

The students were given no special guidelines, apart from those provided on the website. The initial study of the tool (Karoulis, Demetriades, & Pombortsis, 2004) showed that it was usable and its use quickly became transparent, so no extra preparations had to be made at that point. The results of this study, concerning the ease of use of the sheet, confirmed those of the aforementioned study.

Measurement Instruments

The main means of depicting the students' performance was the final examination, which consisted of 15 questions, covering all aspects discussed in the pdf files. An extra percentage scale, from 0 percent to 100 percent, was added at the header of the document, where the students pinpointed the usage of the lesson sheet during their study.

The questions were of quantitative type in order to be easily analyzed statistically. Five kinds were employed in this experiment:

- Multiple choice: one of several correct answers is possible
- Yes/no or right/wrong: bipolar differentiated notions
- Choose from a list: more than one can be correct
- Matching: pairing corresponding items from two columns
- Fill in the blank: the student must fill in the correct word

Appendix B shows a sample of the examination sheet used (one page out of four) presenting some of the questions used.

Experiment

Preparation

The preparation of the experiment underwent the following steps, each corresponding to the experimental components:

- Subjects. The students were asked to express their interest in participating in the experiment by having personal contact with one of the tutors or communication via e-mail.
- Material preparation. The lesson framework (what to teach), was defined; the methodology (how to teach) was chosen; the lesson sheet and other files (by what means) were prepared prior to the series of lessons; and the support site (for use during the lessons) was set up.
- Instruments. What had to be examined was decided. Once the pool of questions had been defined, the questions to appear on the examination sheet were chosen.

The procedure described here took one and a half months. Only after the final test were the subjects informed in detail via the website about the experiment in which they took part and its basic structure, yet they knew from the beginning that research would also, in parallel, take place.

Execution

The execution of the experiment was as predicted. There was no deviation from the plan. All students downloaded the study files and kept the schedule. It must be emphasized at this point, that the final test was performed under real circumstances, in separate rooms with the aid of three additional overseers.

Finally, a statistical analysis, as subsequently described, was applied in order to investigate any significant difference in the performance of the two student groups.

Data validation

The data collected through the examination sheets was considered reasonable, as there was no indication to the contrary. Data was also considered to have been collected correctly, as it was recorded on the examination sheets, and the subjects were considered serious (applied correct treatments in correct order), as they were well motivated.

Data analysis & interpretation

After the separation of the students into both groups, the statistical analysis of their performance was made. Descriptive statistics showed that the mean values of the two groups already provide an obvious differentiation, 2.84 to 5.39, nearly double. Standard deviation had a relative low value in both cases (1.23 and 1.34), which means that the students' grades were more or less concentrated around the mean values. Table 1 provides these facts.

Table 1. Descriptive statistics of control and treatment group

Group Statistics				
Use of the tool	N	Mean	Std. Deviation	Std. Error Mean
No (< 50%)	60	2.8442	1.23296	.15917
Yes (> 50%)	36	5.3903	1.34717	.22453

Ranks				
Use of the tool	N	Mean Rank	Sum of Ranks	
Grade No (< 50%)	60	33.60	2016.00	
Grade Yes (> 50%)	36	73.33	2640.00	
Total	96			

The data of table 1 is depicted graphically in the box-plot and bar graphs in Figure 1.

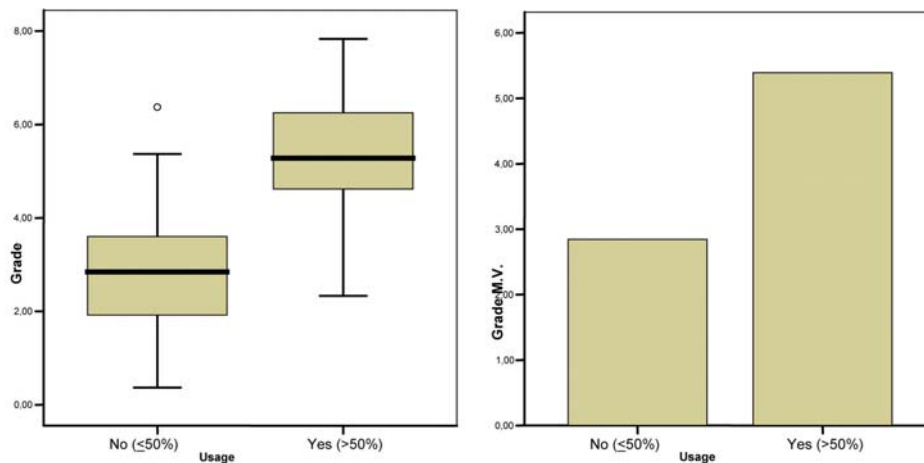


Figure 1. Box-plot (a) and bar graph (b) of the mean values of the grades of the students

This statistically significant differentiation, visible in these graphs, had to be investigated with a series of inferential statistical tests, which are subsequently presented.

Initially, a *t*-test showed a statistically significant differentiation in a statistical significance level of $p < 0.0005$, which is unusually low, as shown in Table 2 below.

Table 2. Results of the t-test
Test of independent samples

	Levene's test for equality of variances		t-test for equality of means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	.256	.608	-9.460	94.000	.0005	-2.54611	.26915	-3.08051	-2.01171
Equal variances not assumed			-9.251	68.724	.0005	-2.54611	.27523	-3.09521	-1.99701

A comparable result also came from the non-parametric Mann-Whitney test (significance level $p < 0.0005$) as well as the Wilcoxon-W test, which were both applied to strengthen this result.

Table 3. Non-parametric Mann-Whitney and Wilcoxon W tests
Test statistics ^a

	Grade
Mann-Whitney	186.000
Wilcoxon W	2016.000
Z	-6.766
Asym. Sig. (2-tailed)	.0005

^a Grouping variable: tool usage

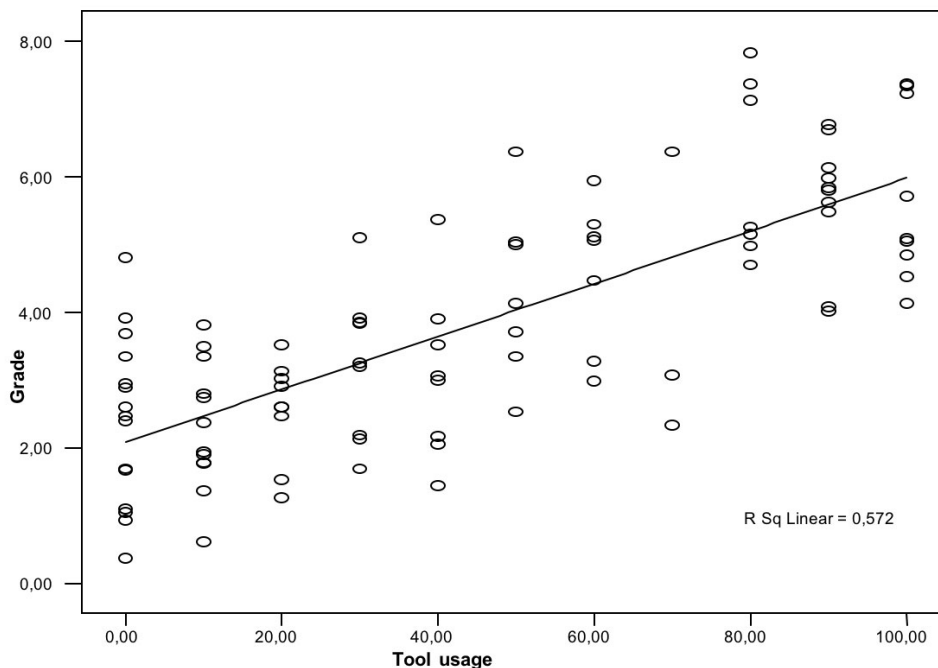


Figure 2. Scatterplot of correlation between grade and usage of the tool

A new question emerged here, namely, whether the usage of the lesson sheet had an impact on the final grade of the student. In other words, did frequent use of the lesson sheet improve the student's score on the final test? This statement obviously enhances the stated research hypotheses; however, it was not initially set.

Because of this, the researchers had to go a step further. First, a dispersion scatterplot was calculated, shown in Figure 2. A linear correlation is apparent. To prove this claim, the Pearson-*r* correlation coefficient has been employed.

Table 4, shows that the Pearson-*r* (0.756) is statistically significant at the 0.01 level and the determinant coefficient ($r^2 = 0.572$) shows that a significant part of the dispersion of the grades (57.2%, shown in Figure 2) can be statistically explained as caused by the use of the lesson sheet.

Table 4. Pearson-r correlation coefficient

		Correlations	
		% Tool usage	Grade
% tool usage	Pearson Correlation	1	.756**
	Sig. (2-tailed)	.	.000
	N	96	96
Grade	Pearson Correlation	.756**	1
	Sig. (2-tailed)	.000	.
	N	96	96

** . Correlation is significant at the 0.01 level (2-tailed)

Interpretation

It is obvious that the group that claimed to have used the tool more than 50 percent performed significantly better than the other group. This result is clearly depicted in the box-plots.

Hypotheses answering

So, the stated hypotheses can now be answered. According to what has already been discussed, the null hypothesis H_0 must be rejected and the alternative H_a adopted instead. Use of the lesson sheets does affect students' performance on the test.

However, this difference could always be the result of a third factor causing this effect (causal relation). In other words, the "good" students could very well be the ones that really used the tool and benefited from its impact. This is a serious consideration and will be discussed later.

Generalization

Generalization can only be made in the context set by the particular experiment. However, Wohlin et al. (2000, p. 112–113) provide a good discussion on conclusions that can be drawn and generalizations that can be made in a broader context.

To make an attempt for any generalization, one must consider the external validity of the experimental settings. As discussed in the next section, the experiment was performed in real circumstances, so subjects, settings, and history do not threaten generalization. Therefore, it can be argued that the results of this study can be generalized in the context of university-level lessons. However, there is a concern regarding the cognitive domain to be taught, so this generalization must be considered as tentative, as the particular cognitive domain is an important factor.

Practical importance

The practical importance of the results presented here could be of interest to anyone dealing with ODL. Educational designers especially are seeking any aid in the direction to support, scaffold, and guide the distance-education student, so this proposed tool, even in its printed form, could provide a fair solution, as shown by the studies so far.

As for a potential replication of the experiment, the described context can be easily mimicked, so a replication of this experiment is always possible.

Validity evaluation

The validity of the conclusions is highly dependent on how well specific threats were manipulated. According to Cook and Campbell (1979), validity is subdivided into four types, each of which addresses a specific methodological question and includes one type of threat to the validity of the results. The four threats are conclusion, internal, construct, and external validity. There is a strong relationship between these four validity threats (Trochim, 1999; Wohlin, 2000, p. 64), depicted in Figure 3.

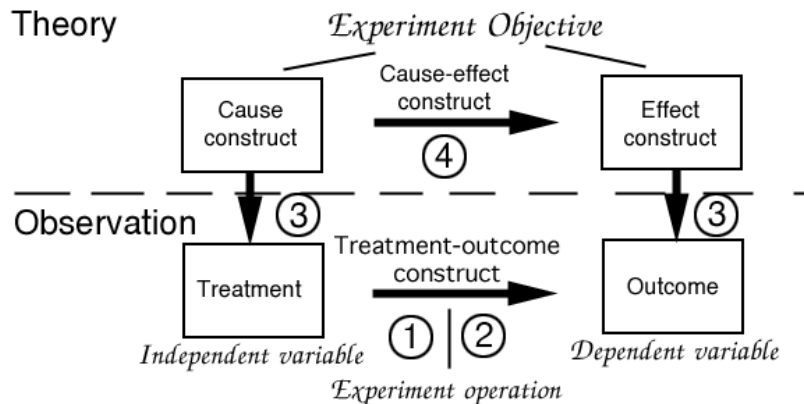


Figure 3. Relationships in the experiment design with respect to validity. Adapted from Trochim (1999)

The main idea behind this relationship is that, in order to draw valid conclusions about the theory defined in the hypotheses, based on observations, one must provide satisfactory answers to all the specific methodological questions that evolve. So, in this section, the four validity threats are discussed as well as the approaches employed to confront them. However, aspects that are irrelevant to the particular context, or obviously not applicable, were neglected.

Conclusion validity

Conclusion validity, point 1 in Figure 3, is concerned with ensuring that there is a statistical relationship between treatment and outcome. So, threats to conclusion validity are factors that can lead to incorrect conclusions about relations between the treatment and the outcome of the experiment. According to Cook and Campbell (1979), factors such as low statistical power or violated assumption of statistical tests threaten conclusion validity. In this case, the reliability of the measures is ensured by the use of the examination sheets. So, as this approach constitutes an objective rather than subjective measurement mode, only unintentional errors could occur during the grading phase. Because the experiment occurred under real circumstances, during a university-level lesson, the reliability of treatment implementation was also ensured, as all subjects were treated in a similar manner.

Internal validity

Internal validity, point 2 in Figure 3, means that the observed relationship is a result of the factor (i.e., causal). Threats to internal validity are influences that can affect the independent variable with respect to causality, that is, they can affect the conclusion of a possible causal relationship between treatment and outcome. According to Cook and Campbell (1979), factors such as fatigue effects (negative), learning effects (positive), or instrumentation may threaten internal validity. Although single group threats of internal validity are considered to affect only experiments with no control group to which the treatment is applied (Wohlin et al., 2000, p. 69), some of these factors may affect this experiment as well. However, the quantitative approach followed to assess the performance of

the students diminishes the threat of an eventually suboptimal instrumentation. Regarding the selection of subjects, volunteers are considered to be generally more motivated (Wohlin et al., 2000, p. 69), as it was in this case.

Construct validity

Construct validity, point 3 in Figure 3, concerns the generalization of the experimental result to concept or theory behind the experiment. So, if the relationship between cause and effect is causal, then treatment reflects cause and outcome reflects effect.

Inadequate pre-operational explication of constructs: The main concern in the described experiment is whether an examination sheet really depicts the acquisition of offered information, and what exactly is measured by the examination sheet. To confront these aspects, the examination sheets have been designed according to two factors. Firstly they completely covered the taught domain: seven questions concerned the first lesson (usability engineering — design) and eight questions concerned the second lesson (usability engineering — evaluation). So it can be argued that the acquisition of the offered information must be complete and broad on the domain, in order for the subject to perform well on the test. Secondly, the quantitative questions aim to measure this acquisition in a clear scale from 1 to 10 for every question, giving a total of 150 points. So, it can be argued that the acquisition of the information can also be measured, yet the assimilation of the information cannot. The passing from memorization to learning and understanding is a complicated issue, and was outside the scope of this study. However, it is discussed in Karoulis, Demetriades, & Pombortsis (2004). To conclude, the described design of the examination sheet aimed to confront this specific threat as well.

Mono-operation and mono-method bias: As the experiment concerned only one domain (usability engineering) and used only one measurement instrument (the examination sheet), this threat was present. The strict design and conduct of the experiment diminishes these threats.

External validity

If there is a causal relationship between cause and effect, can it be generalized? Therefore, external validity, point 4 in Figure 3, is related to generalizing. By making the experimental environment as realistic as possible, threats to external validity are reduced and research results can be both generalized to the population under study as well as to other research settings.

According to Cook and Campbell (1979), interactions between selection, settings, history, and treatment may threaten external validity. However, in this case the real settings of the experiment diminish most of those threats as well.

Discussion and concerns

However, the greatest threat to the presented validity evaluation was omitted, on purpose. It will now be discussed in detail. One point in Conclusion Validity (random heterogeneity of subjects) and one point in Internal Validity (selection and randomization) cannot be considered as addressed by the experiment design and remain as validity threats. In more detail, a very plausible assumption could well be that only the “good” students adhered to the guidelines provided by the lesson sheet and really used the tool as it was proposed; therefore, it is a reasonable conclusion to assume that they performed better in the test. The results from the first study on the lesson sheets, where only the good students accepted the tool, while the rest ignored it, strengthen this claim.

However, there are some reasons to believe that this is not as bad as it seems. Firstly, education in general underlies this threat, so why not ODL? In every school a number of pupils and students try to cheat, do not follow instructions and are accordingly characterized as “bad.” This is a social phenomenon that is completely outside of the scope of this study.

Secondly, ODL presumes the personal responsibility of the, usually adult, student. The inactivity, loss of interest, and finally dropout phenomenon is a well-known problem in ODL. This is a main concern in this study. The distance

student needs support, guidance, and scaffolding, elements that usually are not present in the instructional material. Many studies up to now, such as Bernath and Rubin (2004), Pierrakeas, Xenos, Panagiotakopoulos, and Vergidis, (2004), and Xenos, Pierrakeas, and Pintelas (2002), pinpoint that the drop-out rates vary between 25 percent and 30 percent, in a roughly estimated mean value, as it strongly depends on the course and educational context. From this point of view, the encouraging results of the present study argue for the use of the tool in more cases, because only a broad application and a number of reports on its effectiveness can clarify its potency.

Another concern of the present study, not visible in the validity evaluation, is the construction of the educational environment. Contemporary distance education separates the instructor from the students in time and space, yet provides remarkable educational results outside the context of the traditional class (Keegan, 1996). It is also argued that the use of enriched multimedia material provides a better instructional environment and augments the quality of the delivered education (Owoc, Maciaszek, & Hauke, 2001). In our case, these attributes were present, specifically the printed educational material combined with the multimedia presentation. However, it is debatable if such a poor educational setting constitutes a real ODL environment. There were, however, some specific reasons for this approach. Firstly, the validity of the experiment had to be ensured, so a very simplified ODL environment was favored over a more innovative one. Although this reason could be enough, as it confronts many validity questions, it was also argued that the potency of a new proposed tool should also be examined in an environment familiar to the subjects, to avoid any special adaptation training that could interact with the experiment settings. Such a “minimalist” approach has also been reported to perform adequately (Karoulis, Sfetsos, Stamelos, Angelis, & Pombortsis, 2004), so the present study could focus untrammelled on the effectiveness of the lesson sheets.

However, despite these claims, the question of the validity of the results of this study in a real ODL environment (with tutors, tutorial sessions, delivered over the Internet, etc.) remains unanswered. It is obvious that more research has to be undertaken on this aspect. However, an aid comes through Wedemeyer (1973, p. 101), who defines the notion of “freedom of the distant student” according to three parameters:

1. The learner must be able to define his own pace of learning
2. Learning must be individual and personalisable
3. The learner must be able to choose his instructional targets as well as the means to achieve them.

The point of discussion here is in how far the lesson sheet can assist the distant student in these directions of ODL, both in the presented study, as well as in a more traditional ODL setting. It can be argued that this tool, which can be ad-hoc constructed for every educational context, adheres to these pillars. The basic construction guidelines, presented in section 2, more or less support all three ODL principles stated by Wedemeyer (1973). Former studies, as well as this study, strengthen this claim. However, there are always doubts left, due to the variability of the ODL settings, and the diversity of the entities employed and interacting during any instructional process.

Another concern affects the practical importance of the proposed tool, as already discussed, and its application independent of the cognitive domain. Up to now, the tool has been applied in six different settings. Four were at the post-graduate level and six at the under-graduate level. Although a generalization can hardly be attempted, indications show a common acceptance of the tool by all classes, especially according to qualitative approaches such as interviews. Therefore, more study on multiple domains is needed to draw a valid conclusion regarding broad application of the proposed tool.

Further research

Further research should also consider the transition of the tool to a more technology-based format, such as an interface agent or a browser plug-in. URL resources, animated diagrams, or adaptive interactions with the student are simple possibilities of this approach. The necessary alterations should be investigated, and some formal testing of the adapted tool will provide valuable information on its use in sensitive aspects, such as the scaffolding and support for students who are studying at a distance. From a technological perspective, if the tool is successfully implemented in ODL, it could greatly aid in solving the problem student dropout, which is the number one threat in almost all ODL programs.

Conclusion

The underlying pedagogy of the lesson sheets has already been published in three prior works. This paper deals only with its evaluation in a distance-learning setting and proposes a more rigorous methodology to do this, based on an extensive statistical elaboration. Although statistical findings do not add anything from a pedagogical perspective in the study, they prove better performance of the group that used the tool.

Based on the aforementioned results, it can be concluded that the lesson sheet, in its present form, can provide a valuable instructional aid in cases of simple ODL environments and an indication of its effectiveness in more complex environments as well.

This conclusion should be kept in mind when it comes to the design of the sheets for application during the instructional procedure. However, it is of paramount importance to always keep in mind the basic construction guidelines, presented in section 2, and adapt them according to the particular educational context.

References

- Basili, V. & Rombach, H. (1988). The TAME project: Towards improvement-oriented software environments. *IEEE Transactions on Software Engineering*, 14.
- Basili, V. & Weiss, D. (1984). A methodology for collecting valid software engineering data. *IEEE Transactions on Software Engineering*, 10.
- Belanger, F. & Jordan, D. H. (2000). *Evaluation and Implementation of Distance Learning: Technologies, Tools and Techniques*. Hershey, PA: IDEA Group.
- Bernath, U. & Rubin, E. (2004). Student retention issues in the online master of distance education (MDE): An evidence-based approach. In U. Bernath & A. Szücs (Eds.). *Supporting the Learner in Distance Education and E-Learning, Proceedings of the Third EDEN Research Workshop* (pp. 363–369). Carl von Ossietzky University of Oldenburg, Germany.
- Cook, T. D. & Campbell, D. T. (1979). *Quasi-Experimentation—Design and Analysis Issues for Field Settings*. Boston, MA: Houghton Mifflin Company.
- Forsyth, D. R. and McMillan, J. H. (1991). Practical proposals for motivating students. *New Directions for Teaching and Learning, Vol 1991* (45), 53–65.
- Georgiadou, E. & Economides, A. (2000) Evaluation Factors of Educational Software. *Proceedings of IWALT 2000 Conference* (pp. 113–116). Palmerston North, New Zealand.
- Karoulis, A. Demetriades, S. & Pombortsis, A. (2004). The “lesson sheets” and their application in the instructional procedure. *IEEE Educational Technology & Society*, 7(2), 105–119.
- Karoulis, A., Stamelos, I., Angelis, L., & Pombortsis, A. (2004). Formally assessing an instructional tool: A controlled experiment in software engineering. *IEEE Transactions on Education*, 48(1), 133–139.
- Karoulis, A., Sfetsos, P., Stamelos, I., Angelis, L., & Pombortsis, A. (2004). Experimentally assessing a resource-effective design for ODL environments. *Interactive Technology and Smart Education Journal*, 1(4), 265–279.
- Keegan, D. (1996). *Foundations of distance education* (3rd ed.). London: Routledge.
- Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice Hall.
- Lewis, C. & Rieman, J. (1994). *Task-centered user interface design: A practical introduction*. Retrieved September 19, 2003, from <ftp://ftp.cs.colorado.edu/pub/cs/distrib/clewis/HCI-Design-Book/>
- Lowman, J. (1981). *Mastering the techniques of teaching*. San Francisco: Jossey-Bass.

Owoc, M. L., Maciaszek, L. A., & Hauke K. (2001). On principles of course evaluation in distance learning environment. *Proc. of Informing Science Conference 2001. Challenges to Informing Clients: A Transdisciplinary Approach*. Krakow, Poland, June 19-22, pp. 385-391.

Pierrakeas, Ch., Xenos, M., Panagiotakopoulos, Ch., & Vergidis, D. (2004). A comparative study of dropout rates and causes for two different distance education courses. *International Review of Research in Open and Distance Learning*, 5(2).

Squires, D. & Preece, J. (1996). Usability and learning: Evaluating the potential of educational software. *Computers and Education*, 27(1), 15–22.

Trochim, W. (1999). *The research methods knowledge base* (2nd ed). Ithaca, NY: Cornell.

van Houten, R. (1980). *Learning through feedback: A systematic approach for improving academic performance*. New York, NY: Human Sciences.

Xenos, M., Pierrakeas, C. & Pintelas, P. (2002). A survey on student dropout rates and dropout causes concerning the students in the course of informatics of the Hellenic Open University, *Computers & Education*, 39(4), 361–377.

Wedemeyer, C. (1973). The use of correspondence education for post-secretary education. In A. Kabwasa & M. Kaunda (Eds.). *Correspondence education in Africa*, London: Routledge & Kegan.

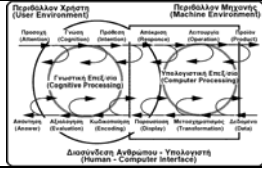
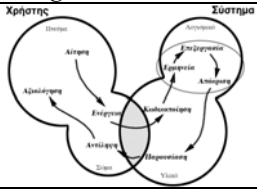
Wohlin, C., Runeson, P., Höst, M., Ohlsson, M., Regnell, B., & Wesslén, A. (2000) *Experimentation in software engineering*. Norwell, MA: Kluwer.

Appendix A: A sample lesson sheet

Learning Guide

Use this learning guide as a support to your study. Adhere to the flow of this guide and follow the instructions. You will use the learning material that you have downloaded. It is important to elaborate the proposed exercises and to work with the additional material, as pinpointed by this guide. You can make notes on this sheet, thus personalizing it to your educational demands. It will not be returned to the instructor but will become a study source and reference for your own use.

Part A: Usability Engineering

1. Introduction — the field																	
<p>Interface (pages 19–21 on the HCI01.pdf file that you have downloaded) (Explain the differences — Note the important aspects)</p>	<p>Interface: Interaction: Interfere:</p>																
<p>The picture by Norman K. (page 21) (Study it for a couple of minutes and make comments on it on the sheet)</p>																	
<p>Exercise 1: In a few words, explain the presented terms in the interaction schema above, as well as their contribution to the interaction <i>Write your answer here</i></p>	<p><i>It is not important if you cannot explain all! Full explanations could cover many pages...</i></p>																
<p>Interface as “Inter-face” (page 22) ATTENTION! <i>NOT in your handbook! →</i></p>	<p>Definition <i>Interface is...</i> Other definitions: <ul style="list-style-type: none"> • Preece: the sides of the system... • Shneiderman: </p>																
<p><i>Remember to enrich this sheet with your own comments, notes, suggestions, etc. Your final review will be a passage through this sheet.</i></p>																	
<p>A figure for this definition (page 23 → read your book AFTER you have studied this figure and AFTER you have performed the exercise below.)</p>																	
<p>Exercise 2:</p>	<p>a) Describe aloud the human-system interaction according to this figure (start at “demand”) b) What does the grey area mean? (The answer is at the end of this sheet.)</p>																
2. Man–Machine Interaction and Usability																	
<p>Definitions (page 24)</p>	<p>ACM (1992) defines human–computer interaction as <i>the cognitive domain...</i></p>																
<p>Exercise 3: Mark the corresponding terms <i>Fill in the blanks</i></p>	<table border="1"> <thead> <tr> <th></th> <th>Effectiveness</th> <th>Efficiency</th> <th>Satisfaction</th> </tr> </thead> <tbody> <tr> <td>Learning ease</td> <td></td> <td></td> <td></td> </tr> <tr> <td>Low error rate</td> <td></td> <td></td> <td></td> </tr> <tr> <td>Remembering</td> <td></td> <td></td> <td></td> </tr> </tbody> </table> <p><i>The answer is provided at the end of the sheet</i></p>		Effectiveness	Efficiency	Satisfaction	Learning ease				Low error rate				Remembering			
	Effectiveness	Efficiency	Satisfaction														
Learning ease																	
Low error rate																	
Remembering																	
<p>Additional stuff</p>	<p>Visit the following URL and pinpoint aspects relevant to usability: http://www.a_relevant_url_here.edu</p>																

Appendix B: A Questionnaire sample

Test in Usability Engineering

Name: _____ ImNr: _____ Date: ____/____/ 200____

Score: _____

Questions *(Please note an X only where appropriate)*

1. According to ACM, the fundamental domains of Informatics are:

- | | |
|---|--|
| <input type="checkbox"/> Hardware | <input type="checkbox"/> Communications systems |
| <input type="checkbox"/> Windows | <input type="checkbox"/> Arithmetic and symbolic computing methods |
| <input type="checkbox"/> Artificial intelligence and robotics | <input type="checkbox"/> Logical programming |
| <input type="checkbox"/> Distributed systems | <input type="checkbox"/> C language and UNIX |
| <input type="checkbox"/> Object-oriented programming | <input type="checkbox"/> Operation research |
| <input type="checkbox"/> Databases and information retrieval | <input type="checkbox"/> Digital electronics |
| <input type="checkbox"/> Human-computer interaction | <input type="checkbox"/> Software engineering |
| <input type="checkbox"/> Data structures | |

2. Which one is the correct *Human-computer communicative interaction*?

- Action – Interpretation – Coding – Processing – Display – Perception– Evaluation
- Demand – Coding – Interpretation – Processing – Presentation – Evaluation – Perception
- Demand –Action– Coding – Interpretation – Processing – Feedback – Presentation – Perception – Evaluation
- Coding – Demand – Interpretation – Processing – Presentation – Evaluation – Perception
- Action– Demand – Coding – Interpretation – Processing – Feedback – Presentation – Evaluation – Perception
- Demand – Action – Coding – Action – Processing – Presentation – Perception – Evaluation
- Action – Coding – Interpretation – Processing – Display – Perception – Evaluation – Feedback
- Action – Interpretation – Processing – Coding – Display – Evaluation – Perception – Feedback

3. At the definition of ISO: “*Usability = Effectiveness + Efficiency + Satisfaction*” define:

- | | | | |
|------------------------|--|-----------------------|--|
| <i>Effectiveness</i> = | <input type="checkbox"/> Efficiency + Accuracy | <i>Satisfaction</i> = | <input type="checkbox"/> Positive Feeling + Positive Attitudes |
| | <input type="checkbox"/> Accuracy + Completeness | | <input type="checkbox"/> Positive Feeling / Positive Attitudes |
| | <input type="checkbox"/> Utility + Accuracy | | <input type="checkbox"/> Happiness + Positive Attitudes |
| | <input type="checkbox"/> Efficiency / Resources | | <input type="checkbox"/> Comfortability + Positive Attitudes |
| | <input type="checkbox"/> Validity + Accuracy | | <input type="checkbox"/> Happiness + Positive Feeling |
| | <input type="checkbox"/> Accuracy / Validity | | <input type="checkbox"/> Comfortability + Validity |
| | <input type="checkbox"/> Validity / Accuracy | | <input type="checkbox"/> Utility + Validity |
| | <input type="checkbox"/> Validity / Resources | | <input type="checkbox"/> Positive Feeling + Positive Attitudes |

4. Which are the usability parameters, according to Nielsen?

- | | |
|---|---|
| <input type="checkbox"/> Efficiency of interaction | <input type="checkbox"/> Low user error rate |
| <input type="checkbox"/> Satisfactory feeling | <input type="checkbox"/> Clear file structure |
| <input type="checkbox"/> Subjective user satisfaction | <input type="checkbox"/> Navigation speed |
| <input type="checkbox"/> Error protection | <input type="checkbox"/> Easy to learn |
| <input type="checkbox"/> High work efficiency | <input type="checkbox"/> Easy to remember the use of the system |
| <input type="checkbox"/> High user efficiency | <input type="checkbox"/> Easily recovery from errors |
| <input type="checkbox"/> Efficient work | <input type="checkbox"/> Easy navigation |
| <input type="checkbox"/> Satisfactory system efficiency | |

5. Which of the following is correct for the design models?

- | | |
|---|--|
| <input type="checkbox"/> The waterfall model demands resource definition. | <input type="checkbox"/> The star model functions after the completion of the system |
| <input type="checkbox"/> The waterfall model includes requirement analysis, design, implementation, and evaluation. | <input type="checkbox"/> The star model doesn't include the implementation phase |
| <input type="checkbox"/> The waterfall model is iterative. | <input type="checkbox"/> The spiralette is iterative |
| <input type="checkbox"/> The waterfall model is based on evaluation. | <input type="checkbox"/> The spiralette is user-centered |
| <input type="checkbox"/> The waterfall model demands a working team. | <input type="checkbox"/> The spiralette is based on evaluation |
| <input type="checkbox"/> The star model is based on evaluation. | <input type="checkbox"/> The spiralette demands resource definition. |
| <input type="checkbox"/> The star model demands resource definition. | <input type="checkbox"/> The spiralette is applied after the completion of the system. |
| <input type="checkbox"/> The star model is user-centered. | |