Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Author's personal copy

Available online at www.sciencedirect.com
ScienceDirect
Journal of Visual Languages and Computing 19 (2008) 721–737
Journal of Visual Languages & Computing
www.elsevier.com/locate/jvlc

Assessing the usability of a visual tool for the definition of e-learning processes

Gennaro Costagliola[a], Andrea De Lucia[a], Filomena Ferrucci[a], Carmine Gravino[a], Giuseppe Scanniello[b,*]

[a]*Dipartimento di Matematica e Informatica, Università di Salerno, Via Ponte Don Melillo, 84084 Fisciano, SA, Italy*
[b]*Dipartimento di Matematica e Informatica, Università della Basilicata, Viale dell'Ateneo Macchia Romana, 85100 Potenza, Italy*

Received 17 October 2006; received in revised form 15 November 2007; accepted 28 January 2008

Abstract

In this paper, we present a usability study aiming at assessing a visual language-based tool for developing adaptive e-learning processes. The tool implements the adaptive self-consistent learning object SET (ASCLO-S) visual language, a special case of flow diagrams, to be used by instructional designers to define classes of learners through stereotypes and to specify the more suited adaptive learning process for each class of learners. The usability study is based on the combined use of two techniques: a questionnaire-based survey and an empirical analysis. The survey has been used to achieve feedbacks from the subjects' point of view. In particular, it has been useful to capture the perceived usability of the subjects. The outcomes show that both the proposed visual notation and the system prototype are suitable for instructional designers with or without experience on the computer usage and on tools for defining e-learning processes. This result is further confirmed by the empirical analysis we carried out by analysing the correlation between the effort to develop adaptive e-learning processes and some measures suitable defined for those processes. Indeed, the empirical analysis revealed that the effort required to model e-learning processes is not influenced by the experience of the instructional designer with the use of e-learning tools, but it only depends on the size of the developed process.
© 2008 Elsevier Ltd. All rights reserved.

Keywords: Adaptive learning processes; e-Learning; Effort estimation; Usability survey; Visual languages

1. Introduction

The academic and industrial realities are proposing several software tools supporting instructional designers during the development of e-learning processes delivered via the Web. Instructional designers have to consider the structure of an e-learning process in terms of didactic contents, dependences, and assessment rules. In general, the definition and the development of e-learning processes are influenced by teaching domain, knowledge presentation, and multimedia objects enhancing the didactic contents [1–4].

Visual and diagrammatic representations have been introduced to support instructional designers during the development of learning processes [5–7]. As a matter of fact, visual languages play a central role in several application domains as they are

---

*Corresponding author. Tel.: +39 3204221575.
E-mail addresses: gcostagliola@unisa.it (G. Costagliola), adelucia@unisa.it (A. De Lucia), fferrucci@unisa.it (F. Ferrucci), gravino@unisa.it (C. Gravino), giuseppe.scanniello@unibas.it (G. Scanniello).

1045-926X/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jvlc.2008.01.003

Author's personal copy

recognised to be important means for describing and reasoning. Their employment generally improves productivity of expert and non-expert users, as they are easier to learn comparing to textual languages [8,9]. However, this is not always true and the usefulness of visual languages in a specific domain needs to be assessed and verified [10]. Many heuristics, guidelines, and techniques are provided in the literature to assess the usability of software systems and visual language-based tools [8–16].

In [17], the authors have presented the adaptive self-consistent learning object SET (ASCLO-S) visual language, a special case of flow diagrams that can be used by instructional designers to define adaptive e-learning processes. This language has been implemented in a visual-based tool, named ASCLO-S editor. The visual formalism and the tool have been developed within an e-learning research project carried out at University of Salerno aiming at providing teachers and learners with an adaptive context, which enables the development of didactic contents considering self-expectations, satisfaction, welfare, and diversities among learners. Let us note that we decided of developing ASCLO-S editor since commercial and academic tools providing the needed functionalities were lacking. Successively, the effectiveness of the ASCLO-S editor has been assessed by comparing the tool with Macromedia Authorware [18], a widely employed commercial authoring tool. Then, before starting the experimentation phase of the tool in the academic context, it was crucial to carry out a usability study meant to evaluate the efficacy of the ASCLO-S visual language and the usefulness of the supporting tool. In the present paper, we report on the results of this usability study that is based on the combined use of two techniques, namely a questionnaire-based survey and an empirical analysis, and has been carried out with a group of ten volunteers having heterogeneous background and different teaching experiences.

The survey has been used to achieve feedbacks on the tool from the subjects' point of view. It is based on a questionnaire defined to capture usability concerns of the visual notation and of the graphical user interfaces of the ASCLO-S editor. To this end, the questionnaire is composed of five categories: *subject experience*, *general evaluation*, *special judgement*, *tool learning and usability*, and *information grant*. The analysis of the questionnaire has revealed that the subjects have not manifested any kind of listlessness and indifference but rather a good satisfaction degree and the time spent to learn and use the tool has been considered appropriate. Moreover, the proposed visual notation and the system prototype are considered suitable by instructional designers with or without experience on the computer usage and on tools for defining e-learning processes. This result has been further confirmed by the empirical analysis we have carried out by analysing the correlation between the effort to develop adaptive e-learning processes and some measures suitably defined for those processes, by exploiting an ordinary least-squares (OLS) regression analysis [19]. The aim has been to identify the factors that influence the effort to define and to develop adaptive e-learning processes by using the ASCLO-S editor. The empirical analysis has revealed that the effort to model e-learning processes depends on the size of the developed process, but it is not influenced by the experience of the instructional designer with the use of e-learning tools. It is worth pointing out that in the context of effort estimation, OLS has been widely employed with interesting results [20–22], while this technique has not been adopted in the context of usability studies.

The remainder of the paper is organised as follows. Section 2 presents related work. Section 3 describes the ASCLO-S visual language and the tool implementing it, while the design of the usability study is presented in Section 4. The results of the survey and the empirical analysis are discussed in Section 5. Final remarks and future work conclude the paper.

2. Related work

In this section, we discuss related work, in particular, other methods and tools proposed to support the work of instructional designers, and the approaches and empirical studies proposed in the literature to assess the usability of e-learning systems.

2.1. Methods and tools to support instructional designers

Muraida and Spector [23] assert that there is "a lack of instructional designer expertise, pressure for increased productivity of designers, and the need to standardise products and ensure the effectiveness of product". Thus, tools supporting instruction design during all the phases of the learning process definition are desirable.

Author's personal copy

Goodyear [24] views the instructional design as falling within four main approaches. These approaches allow the instructional designer to generate e-learning activities from given specifications by means of tools supporting the design of course structure, the selection of presentation templates, the reuse of design elements, and the coordination of activities accomplished by a design team. The ASCLO-S editor effectively supports the first three activities identified by Goodyear. In particular, the analysed editor places great emphasis on the course structure design and the reuse of predefined presentation templates and previously designed didactic activities. On the other hand, the coordination of an instructional design team is a management activity that is typically not supported by authoring tools. Goodyear [25] also proposes an approach for analysing and designing distance courses that is divided into neat parts. The first part of Goodyear's approach resembles the work of other people (outside education) who are interested in the design of technology supporting the work of information system designers, requirements engineers, human factors specialists, and so on. The second part is instead focused on the design of good learning tasks exploiting traditional analysis and design processes. Often, these tools are not able to compensate for the lack of expertise of instructional designers. Vrasidas [4] presents a system to develop hypermedia as part of courses and learning environments delivered on the World Wide Web. The proposed approach details the structuring of information, branching and interactivity, user interface, and navigation through Web-based distance courses.

Differently from the approaches above, the AIMS Project [26] describes a theoretical framework in which the knowledge domain editing and the course editing are distinguished. Differently, from the approach supported by the ASCLO-S editor, the instructional designer first constructs the domain model in term of concepts and links and then the didactic contents are embedded. Cloete and Kotzé in [27] discuss the design of an authoring tool to create interactive didactic contents. In order to support non-technologist instructional designers, this tool is developed considering the usability as the most important requirement. This tool does not integrate a visual environment to design learning process and diversities among learners are not considered in the creation of interactive contents.

2.2. *Assessing the usability of e-learning systems*

Several usability evaluation techniques and guidelines are proposed in the literature aiming at planning and realise usability studies [11–16]. For examples, Ricks and Arnoldy [15] assert that all the usability studies follow the same basic steps although a wide variation of products can be analysed (e.g. software applications, printers, web sites, etc.). Generally, usability produces several benefits [13,16] ranging from the reduction of the training costs to the improvement of the user satisfaction.

A crucial aspect for the success of e-learning projects concerns the satisfaction degree that the people perceive on the used e-learning systems. Indeed, the less the system is usable, the more the learner or the instructional designer spends to master its functionalities [12]. Of course, it is necessary to define usability attributes for e-learning systems [1]. To do that it is necessary to dwell upon the difference between learning management systems (LMSs) and authoring tools. An authoring tool is a software system conceived to define educational materials and/or e-learning processes that are successively deployed in LMSs. LMSs are integrated learning environments conceived to deploy and enjoy educational materials. This scenario highlights that LMSs, authoring tools, and consequently educational materials should be independently analysed, since for each of them different attributes are considered as relevant. In the context of usability of e-learning systems, some studies are devoted to assess LMSs, authoring tools, and didactic contents [1,28–31]. Only marginally the usability of authoring tools is investigated also because these tools are often developed for instructional designers with specific knowledge of e-learning systems.

Concerning the empirical studies, in the last years many researches have been focussing on the estimation of software development cost [32–35], and different techniques have been employed and compared with build prediction models [36–40]. These models are based on measures to gather information about relevant aspects of software products and then manage their development process. The effectiveness of several metrics introduced to predict the effort needed to design and implement software systems has been assessed in a lot of empirical studies [20,22,32–41]. On the other hand, only few empirical analysis using

Author's personal copy

consolidated modelling techniques are proposed to assess the effort to develop didactic contents. Kapp [42] proposes four strategies to obtain an accurate estimate based on similar projects, formulas, bottom up calculation, and industry standard. Bodnar and Henry [43] describe the creation and the validation of a suite of software metrics for multimedia and didactic contents using metrics for problem-oriented languages. In particular, these metrics are applied to software projects developed using Macromedia Authorware [18]. The authors first interview the developers to point out the metrics regarding the developed didactic contents, and then the developed multimedia products are subjectively validated by interviewing the developers responsible for them. These proposals lack of an empirical validation of the proposed metrics.

3. ASCLO-S editor

Distance courses are usually intended for heterogeneous groups of learners, who are often supported only through asynchronous communication. Thus, a personalisation of didactic contents should be taken into account during their design since this is a crucial aspect for an effective delivery of knowledge. In this section, we describe the visual notation of the ASCLO-S language [17], which has been conceived to allow instructional designer to define e-learning processes that adapt themselves to the learners' background. The description of a meaningful visual sentence, within the editor implementing the ASCLO-S language (i.e. ASCLO-S editor), concludes this section.

3.1. The visual notation

The ASCLO-S language is a flow diagram extension. Using it, instructional designers can define classes of learners and then specify for each of them the more appropriate learning process. Different levels of abstraction are provided by the language, so that a learning process can be broken down and structured into a hierarchy composed of four granularity levels: *ASCLO-S*, *ASCLO*, *knowledge fragment*, and *raw contents. ASCLO-S* is a sentence of the ASCLO-S language and it is an aggregation of adaptive self-consistent learning objects (ASCLO). In particular, an ASCLO is considered as a logical collection of knowledge fragments and self-assessment tests. A knowledge fragment is composed of raw contents, which are textual or multimedia objects presented in linear way. We define a learning process as one or more knowledge fragments within an ASCLO arranged in such a way that the knowledge is presented to the learners considering him/her knowledge. Indeed, to properly present the knowledge fragments to a given learner the instructional designer identifies classes of learners that could have benefit from the formative offer. The number of classes of learners is also chosen by the instructional designer.

Class of learners are characterised by a symbolic name. Hence, to define an adaptive learning process the instructional designer first associates the identified symbolic names with the defined learning process and then binds learners' stereotypes with the defined symbolic names. More symbolic names can be associated to an incoming or outgoing arrow. Symbolic names include one or more stereotypes and do not have stereotypes in common. The number of stereotypes contained in each class of learners is properly chosen by the instructional designer. The learning process that will be presented to the learner is the one with the symbolic name containing the stereotype more similar to his/her profile.

Learners' stereotypes and learners' profiles are represented as strings. For example, given an ASCLO-S sentence composed by n ASCLOs and let $ASCLOi$ the ith ASCLO, a learner stereotype is defined as follows:

$$LT^{ASCLO1}LT^{ASCLO2}\cdots LT^{ASCLOn}LT^{CG},$$

where the generic LT^{ASCLOi} is the linguistic term representing the knowledge degree on the ith ASCLO. On the other hand, LT^{CG} is the linguistic term representing the prior knowledge degree before taking up the course. Four different linguistic terms can be specified: Nk (no knowledge), Lk (low knowledge), Mk (medium knowledge), and Hk (high knowledge).

Fig. 1 shows the visual tokens of the ASCLO-S language. An ASCLO is a composition of knowledge fragments (Fig. 1A), transition elements (Fig. 1B), self-assessment test (Fig. 1C), and joint/disjoint symbols (Fig. 1D). Transition elements are labelled using the symbolic name of a given learners' class. On the other hand, the self-assessment test symbol has to be placed at the end of each ASCLO to update the learner knowledge profile, on the basis of the test results. The proposed language allows the instructional designer to highlight the ASCLOs using dotted rectangles. Moreover, the instructional designers can arrange ASCLOs that have strictly cohesive contents

Author's personal copy

into swimlanes. The start and the stop markers (Fig. 1E) are used to indicate the initial and final states of each ASCLO. To indicate the initial and final states of an ASCLO-S sentence the start and the stop markers (Fig. 1E) are also used. Further details on the visual formalism can be found in [17].

3.2. The system prototype

An example of a visual sentence within the ASCLO-S editor is shown in Fig. 2. It shows the adaptive learning process related to a part of a software engineering course of the Bachelor program in Computer Science at University of Salerno.

The swimlanes in Fig. 2 propose two different branches depending on the learners' knowledge. For example, in the swimlane on the right-hand side, learners without an appropriate knowledge concerning the software life cycle topic, the knowledge fragment *Modelli e cicli di vita del software*, have to bridge this gap before consuming the knowledge fragment *Analisi e Specifica dei requisiti* that contains raw contents on the requirement analysis.

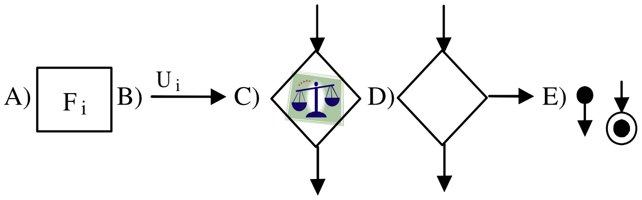


Fig. 1. The visual language icons: (A) knowledge fragment; (B) transition element; (C) self-assessment test; (D) joint/disjoint element; and (E) start and stop marker.

Branches are labelled by the symbolic names: *low_knowledge* and *medium_hign_knowledge*. The association between the symbolic name low_knowledge and the corresponding set of learners' stereotypes is shown in Fig. 3. The user interfaces to define raw contents and questions of a self-assessment test are illustrated in Figs. 4 and 5, respectively. In particular, Fig. 4 shows the raw contents composing the knowledge fragment *UML Introduzione* of the adaptive learning process depicted in Fig. 2. Fig. 5 shows a question belonging to the first self-assessment test of the software engineering course.

4. The design of the usability study

In this section, we present the design of the usability study carried out to assess whether

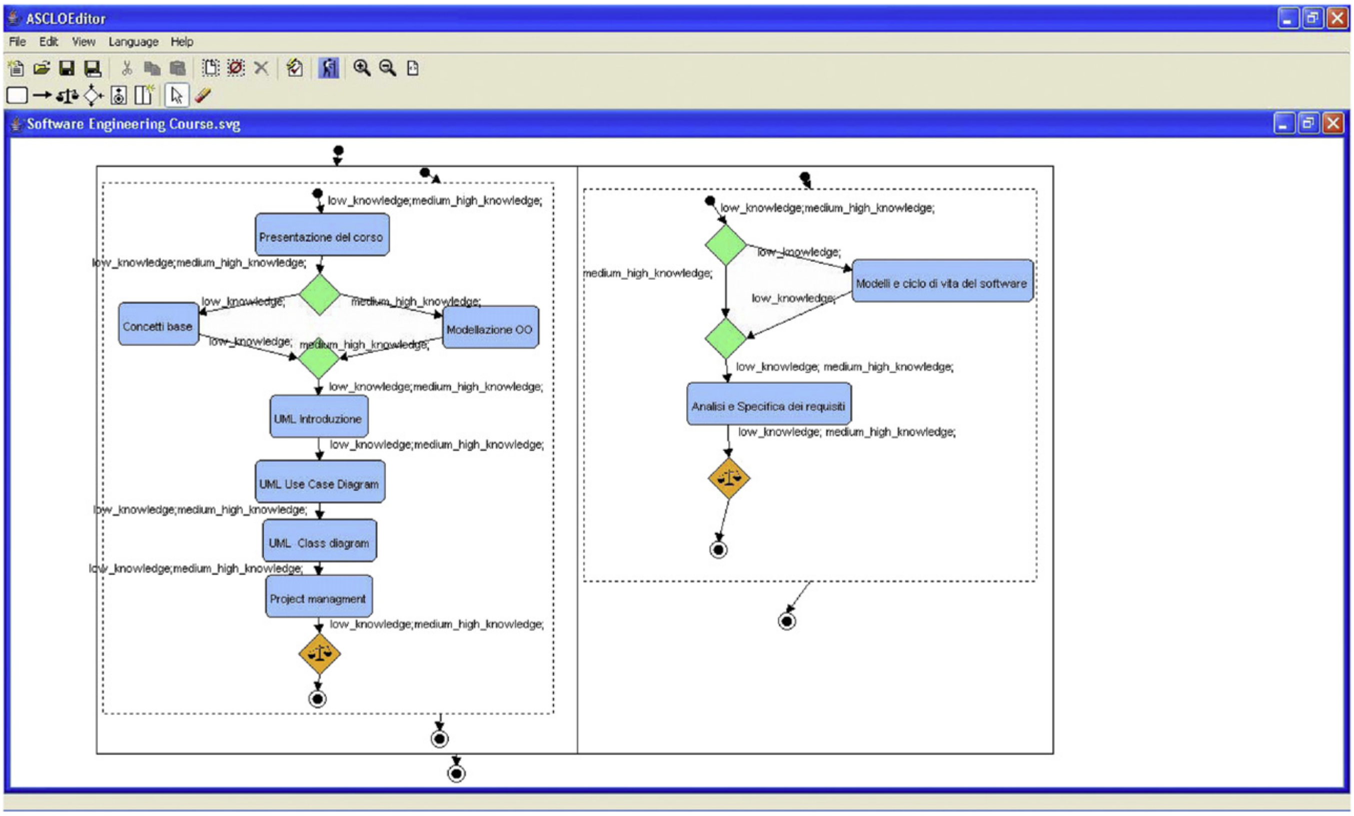


Fig. 2. An adaptive learning process.

Author's personal copy

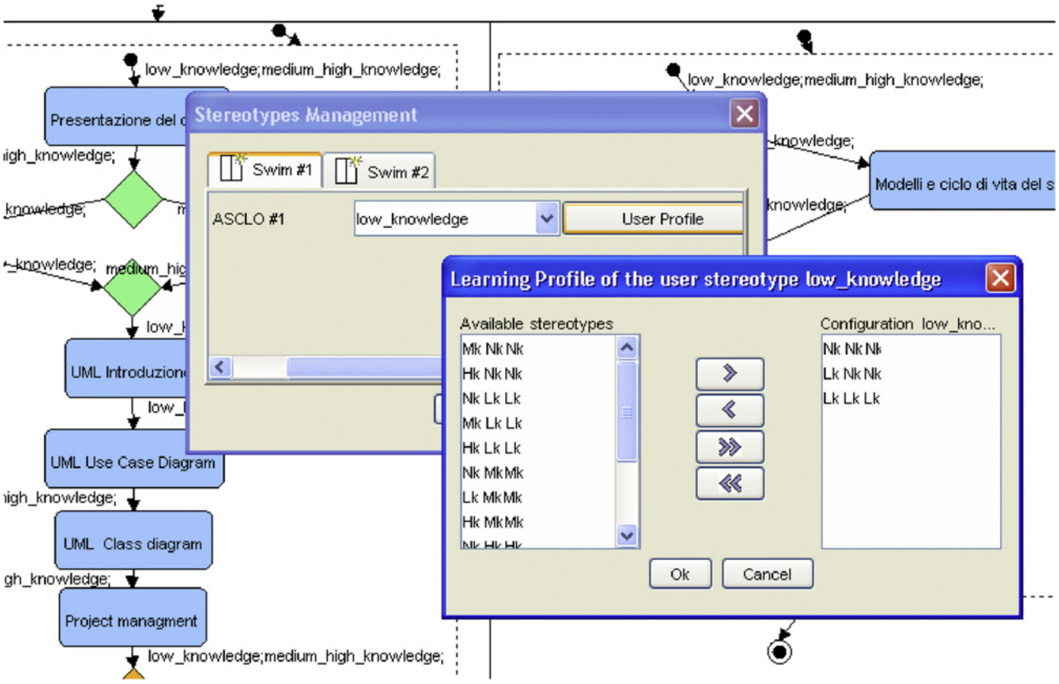
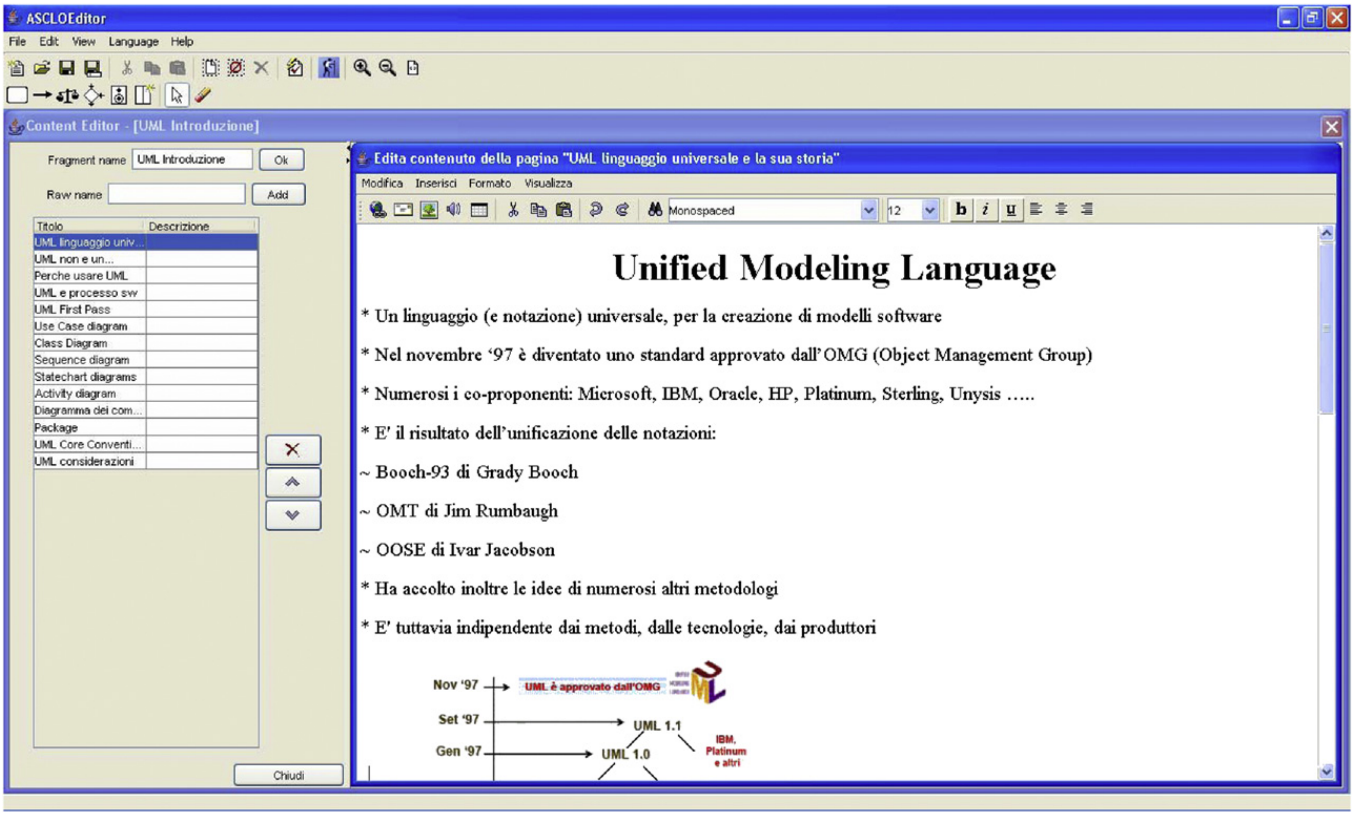


Fig. 3. Associating a symbolic name with the corresponding learner' stereotypes.



Fragment name: UML Introduzione

Raw name:

Titolo	Descrizione
UML linguaggio univ...	
UML non e un...	
Perche usare UML	
UML e processo sw	
UML First Pass	
Use Case diagram	
Class Diagram	
Sequence diagram	
Statechart diagrams	
Activity diagram	
Diagramma dei com...	
Package	
UML Core Conventi...	
UML considerazioni	

Unified Modeling Language

* Un linguaggio (e notazione) universale, per la creazione di modelli software

* Nel novembre '97 è diventato uno standard approvato dall'OMG (Object Management Group)

* Numerosi i co-proponenti: Microsoft, IBM, Oracle, HP, Platinum, Sterling, Unysis .....

* E' il risultato dell'unificazione delle notazioni:

~ Booch-93 di Grady Booch

~ OMT di Jim Rumbaugh

~ OOSE di Ivar Jacobson

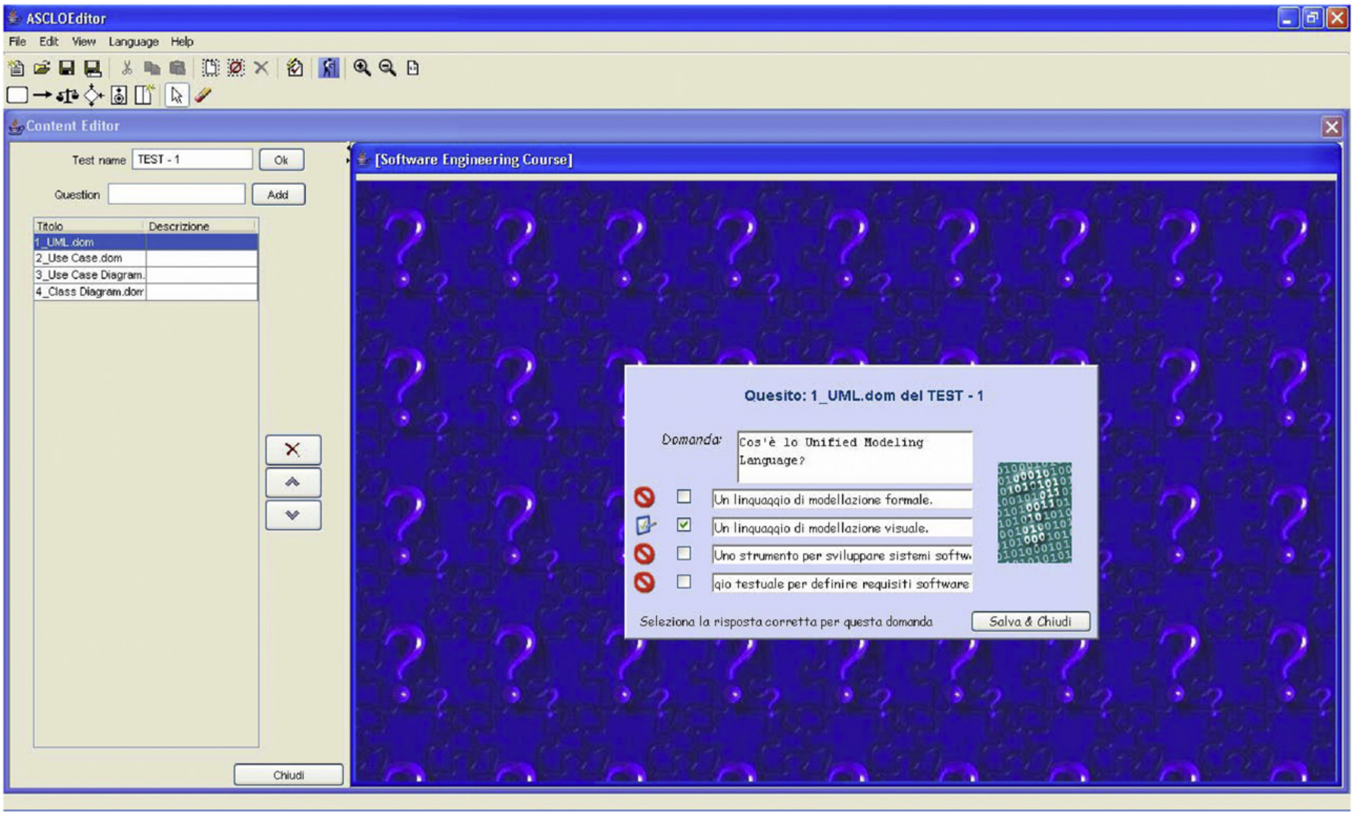
* Ha accolto inoltre le idee di numerosi altri metodologi

* E' tuttavia indipendente dai metodi, dalle tecnologie, dai produttori

Fig. 4. Defining a raw content.

the visual language-based tool effectively supports instructional designers. In the following, we describe the data set and then the techniques we have adopted to assess the usability of the tool as well as the users' subjective satisfaction.

Author's personal copy



Test name: TEST - 1

Question:

Titolo	Descrizione
1_UML.dom	
2_Use Case.dom	
3_Use Case Diagram.	
4_Class Diagram.dom	

Quesito: 1_UML.dom del TEST - 1

Domanda: Cos'è lo Unified Modeling Language?

☐ Un linguaggio di modellazione formale.
☑ Un linguaggio di modellazione visuale.
☐ Uno strumento per sviluppare sistemi softw
☐ gio testuale per definire requisiti software

Seleziona la risposta corretta per questa domanda

Fig. 5. Defining a question of a self-assessment test.

4.1. The data set

Data for the usability study have been gathered considering a group of ten volunteers with heterogeneous background. Regarding the computer literacy, the subjects were able to use the personal computer and productivity programs (e.g. word processors and spreadsheet programs). Table 1 reports on the background of the selected subjects and the corresponding developed e-learning processes.

The usability study has been divided in four steps and performed in one-to-one session (i.e. a supervisor for each subject) using the think aloud technique. All the subjects first have been undergone an introductory course of about a hour on the ASCLO-S editor and its visual notation. Successively, they have been asked to use the tool for 20 min without invoking any kind of tutor support. After that the subjects have performed a task consisting of designing an adaptive e-learning course on a topic on which they had more competence. The subjects have been then asked to deploy the defined adaptive e-learning processes in the E-World framework, an LMS developed at University of Salerno. E-World extends the framework presented in [44] to enable the traceability of adaptive e-learning processes. The traceability of adaptive e-learning processes is managed through a run-time environment compliant with the sharable content object reference model (SCORM) [45] standard of the advanced distributed learning

Table 1
Background and developed e-learning processes

ID	Background	Learning process
1	Student of computer science	Object-oriented software engineering
2	Student of Italian literature	Andy Warhol
3	Teacher of art education	Magna Grecia
4	Manager of a local archaeologist association	Local architecture: "San Pietro a Corte"
5	Sociologist	Pedagogic element for secondary school
6	Teacher of music education	Guitar history
7	Teacher of art education	Ceramic course
8	Ph.D. student in mathematic	Learn to play the guitar
9	Full professor at University of Salerno	XML
10	Research fellow	XSLT

Author's personal copy

(ADL) consortium [46]. Finally, the subjects have filled in a questionnaire to achieve information on tool usability they perceived.

The task had to be carried out respecting some constraints. In particular, they had to design adaptive e-learning processes containing at least two different branches. Moreover, at least one image and a hypertextual link had to be included in the learning process as well as an e-mail address. It is worth noting that, all the didactic material was only available in hard copy. The subjects have also been asked to save and export the created learning process.

4.2. *Design of the survey*

Surveys can be used to collect information on the preferences and reactions of some selected subjects. This technique is particularly valuable for usability investigations since it allows designers to understand the user reaction, to identify possible problems and to enhance the user interfaces. Typically, questionnaires are used to capture the user point of view in order to assess aspects of usability, validity and/or reliability of user interfaces [16]. Questionnaires are usually structured and standardised in order to reduce bias. Questions should be ordered in such a way that each question does not influence the response to subsequent questions and they should be presented in the same order to the subjects.

To collect information regarding the subjects' competences and their perceived usability we have defined a survey questionnaire composed of five categories: *subject experience*, *general evaluation*, *special judgement*, *tool learning and usability*, and *information grant*. In particular, the *subject experience* category is composed of two questions, which aim at knowing the technical competences of the subjects. The general reaction of the subjects in terms of satisfaction degree has been evaluated by the questions belonging to the *general evaluation* category. The questions of the *special judgment* category aim at evaluating the subjects' satisfaction regarding the use of the ASCLO-S tool as well as the perceived usability. The *tool learning and usability* category aims at evaluating the satisfaction degree to master the tool. Finally, the information provided by the tool, while the subject is using it, has been evaluated by the questions of the *information grant* category. The questions of these categories are shown in Table 2.

Table 2
Questions of the usability questionnaire

Category	ID	Question
Subject experience	q11	I use the computer (never, rarely, sometimes, once a week, few days a week, every day)
	q12	My familiarity regarding tools for the definition of learning process and/or didactic contents is (inexistent, inadequate, sufficient, fairly good, good, excellent)
General evaluation	q21	I found the ASCLO-S editor useful
	q22	The usage of the ASCLO-S editor tool is simple
	q23	The feeling with the ASCLO-S editor is satisfactory
	q24	I did not find the ASCLO-S editor boring to use
Special judgement	q31	The user interface is pleasant
	q32	The visual formalism to define adaptive e-learning processes is clear
	q33	The tool proposes clear error messages
	q34	The definition of e-learning processes is effectively supported by the tool
Tool learning and usability	q41	Learning to use the tool is simple
	q42	The required time to use the tool is appropriate
	q43	Remembering the command and their use is appropriate
	q44	The number of steps to perform a task is appropriate
	q45	The effort to insert hyperlinks is appropriate
	q46	The number of steps to insert an image is appropriate
	q47	Learning to define stereotypes is simple
	q48	Learning to export the content is simple
Information grant	q51	Icon names and objects have a clear meaning
	q52	Each set of operations produce a predictable result

The subjects could choose one of the six values reported in Table 2 as answers of questions q1.1 and q1.2, while for the other questions they could specify their judgment within a range. Indeed, these questions expect closed answers according to a Likert scale [47]: from 1 (strongly agree) to 5 (strongly disagree).

Author's personal copy

4.3. Design of the empirical analysis

We have considered the variables reported in Table 3, which have been selected according to three criteria: relevance, easiness to collect, simplicity, and consistency of the counting rules [48]. The first and the second column of Table 3 contain the name of the considered variables and their description, respectively. The type of the considered variables is reported in the third column. The variable *EFH* denotes the effort, expressed in terms of person-hours, to define and create e-learning processes. The remaining variables represent size measures regarding the objects that are specified during the definition of the e-learning processes. Consequently, these measures are determined only analysing the e-learning processes that the recruited subjects defined. It is worth noting that we have employed size measures as our aim has been to verify that the effort required to model e-learning processes depends on the size of the developed process.

In order to graphically represent the distribution of the size measures of the considered e-learning processes, we have adopted the boxplots [49]. They are widely employed in exploratory data analysis since they provide a quick visual representation to summarise the data using five numbers: the median, upper and lower quartiles, minimum (MIN) and maximum (MAX) values, and outliers.

In order to analyse the correlation between the variables denoting the effort to define adaptive e-learning processes and the variables representing information on the processes, we have employed OLS regression analysis. OLS is widely adopted for effort estimation in the context of empirical software engineering [20–22,41]. It is one of the most commonly used statistical techniques for exploring the relationship between a dependent variable and one or more independent variables [19]. The independent variables can be identified by considering those best correlated with the dependent variable, using both Spearman's ρ [50] and Pearson correlations [51]. Indeed, these tests are used to measure the correlation between two variables in terms of a significance level (i.e. p-value).

Table 3
The selected variables

Variable	Description	Scale
EFH	Number of minutes required to develop e-learning processes	Ratio
RC	Number of raw contents	Ratio
Me	Number of media objects	Ratio
ALP	Number of adaptive paths in a learning process	Ratio
KF	Number of knowledge fragments	Ratio
Qs	Number of questions	Ratio
St	Number of possible stereotypes	Ratio

The relationship between dependent and independent variables is described by an equation

$$y = b_1x_1 + b_2x_2 + \cdots + b_nx_n + c,$$

where y is the dependent variable, $x_1, x_2, \ldots, x_n$ are the independent variables, for $i = 1, \ldots, n$, b_i is the coefficient that represents the amount the variable y changes when the variables x changes 1 unit, and c is the intercept.

Some crucial indicators can be used to evaluate the quality of the resulting prediction model. R^2 represents the square of the linear correlation coefficient. This value indicates the amount of the variance of the dependent variable that is explained by the model related to the independent variable. The F-value and the corresponding p-value (denoted by *Sign F*) are useful indicators of the confidence degree of the prediction. A high value of F and a low Sign F suggested that the prediction is indeed possible with a high degree of confidence. We have also determined the p-values and t-values for the model coefficients and intercept. In particular, the p-values give an indication of the accuracy of the coefficient and the intercept, whereas an evaluation of their importance for the generated model is provided by the t-values. A variable is considered significant when its t-value is greater than 1.5, while p-value less than 0.05 are considered an acceptable threshold.

To assess the accuracy of the prediction models, we have taken into account the evaluation criteria: *MMRE* and *Pred(l)* [52]. *MMRE* is the *mean of magnitude of relative error* (*MRE*). *MRE* is defined as

$$MRE = |EFHreal - EFHpred|/EFHreal,$$

where *EFHreal* and *EFHpred* are the actual and the predicted efforts, respectively. According to Conte et al. [52], a value less than 0.25 is an acceptable value for *MMRE. Pred*(l) is the prediction at level l, which is defined as

$$Pred(l) = k/N,$$

where k is the number of observations (Obs), whose *MRE* is less or equal than l, and N is the total Obs.

Author's personal copy

As suggested by Conte et al. [52], a good effort prediction model should have a *Pred*(0.25) value greater than or equal to 0.75.

To obtain non-biased results in the linear regression model evaluation we have carried out a leave-1-out cross validation partitioning at each step the whole data set into a training set and a test set consisting of the remaining observations. The training set is used for model building and the test set for model evaluation. Thus, for each test set we have calculated *MRE*, and then the resulting values have been aggregated across all the sets using *MMRE* and *Pred*(0.25). Indeed, when the accuracy of the model is determined using the same data set used to build the prediction model, the accuracy evaluation is considered optimistic [41]. Leave-1-out cross validation is a widely adopted approach to validate effort estimation models when dealing with small data sets (see, e.g. [20]).

5. Results

In this section, we present the survey and the empirical analysis.

5.1. Survey results

The factors that have been considered relevant for the usability study are time to explore the tool features, time to carry out the usability study, and time spent in inactivity (see Table 4).

The subjects have spent 15 min as mean time, while 6 and 23 min are the best and worst value to explore the tool, respectively. The second column presents the time to accomplish the tasks, which also includes the time to design the adaptive learning process as well as the underlying learning contents. The considerable difference between the best and the worst time to define the e-learning processes depends on its size and the underlying didactic contents. The inactivity time is shown in the third column. Subject3 and Subject8 have been inactive for 1 min, while the remaining subjects have worked without stopping until they had finished their tasks. Finally, Subject5 has been inactive for 2 min.

Table 4
Times of the objective usability findings

Subject	Exploration time	Accomplishment time	Inactivity time
Subject1	10	240	0
Subject2	15	59	0
Subject3	20	45	1
Subject4	18	65	0
Subject5	12	35	2
Subject6	17	58	0
Subject7	15	63	0
Subject8	23	99	1
Subject9	17	61	0
Subject10	6	55	0

The answers of the subjects are summarised in Table 5. This table also reports the mean (*MEAN*) values of the provided answers for each category of the adopted usability questionnaire. In the case of a typical usability study, the adoption of *MEAN* value as descriptive statistic can be considered appropriate to discover as many usability problems as possible. Thus, we have decided to adopt such a descriptive statistic.

The feedbacks obtained by analysing the answers are presented according to the questionnaire categories. The answers of the *general evaluation* category reveal that the subjects have generally manifested a good feeling with the ASCLO-S editor and found it simple and not boring to use. The subjects have also manifested a good satisfaction degree in the definition of the learning process.

The questions of the *special judgement* category reveal that the subjects have manifested a good special judgment on the ASCLO-S editor. All the subjects have found the graphical user interface pleasant except the Subject3, Subject5, and Subject9. In particular, Subject3 and Subject9 have manifested neutral judgment, while Subject5 have found the graphical user interface not pleasant. The majority of the recruited subjects has found the definition of e-learning processes simple and have found the error messages clear and appropriate. Moreover, the totality of the subjects has considered the definition of adaptive e-learning processes simple. We have also gathered relevant suggestions by the subjects concerning the graphical user interface. In particular, Subject3 and Subject5 have found the ASCLO-S Editor user interface lean. Hence, they have suggested of enhancing the user interface with colours to make it more appealing and consequently improve both the perceived usability and the satisfaction degree.

Good results have been also achieved in terms of *Tool Learning and Usability*. The majority of the subjects have found the tool easy to learn. In particular, they have considered the time and the effort to master the tool adequate. The ASCLO-S

Author's personal copy

Table 5
Survey questionnaire results

Questions	Strongly agree (%)	Agree (%)	Neutral (%)	Disagree (%)	Strongly disagree (%)
2.1	20	80	0	0	0
2.2	10	90	0	0	0
2.3	0	100	0	0	0
2.4	0	100	0	0	0
General evaluation	10	90	0	0	0
3.1	0	70	20	10	0
3.2	10	90	0	0	0
3.3	30	70	0	0	0
3.4	0	100	0	0	0
Special judgment	10	90	0	0	0
4.1	10	90	0	0	0
4.2	10	90	0	0	0
4.3	20	80	0	0	0
4.4	30	70	0	0	0
4.5	50	50	0	0	0
4.6	40	60	0	0	0
4.7	10	60	20	10	0
4.8	80	20	0	0	0
Tool learning and usability	30	70	0	0	0
5.1	30	70	0	0	0
5.2	10	90	0	0	0
Information grant	40	60	0	0	0

editor commands have never produced unexpected results and could be easily remembered. The effort to insert multimedia objects as well as the number of steps to accomplish the assigned tasks has been considered more than appropriate by the majority of the subjects. We can also observe that some answers on the stereotype definition simplicity (see question q47) are negative. In particular, 10% of the subjects have found the definition of the learners' stereotypes quite difficult, while 20% have provided a neutral answer. On the other hand, 60% and 10% of the subjects have found the definition of stereotypes quite simple and simple, respectively. The export of a learning process has not produced any problem for almost the totality of the subjects. In particular, 80% and 20% have asserted that the export of the developed learning process is very simple and simple, respectively.

The *information grant* category has revealed that the ASCLO-S editor generally provides a good support. In particular, 40% of the subjects have asserted that the editor provides a good assistance in terms of provided information to use it. The remaining subjects have found this assistance adequate. This appreciable result is due to the fact that all the components of the graphical user interface are understandable and the result of each operation produces always a predictable result.

The observations provided by the subjects have revealed that they have appreciated very much the facilities of the ASCLO-S editor to insert multimedia objects and to export the created knowledge contents. The subjects have also appreciated the results of their work and how the didactic contents have been shown in the Web browser by accessing E-World. Encouraging results in terms of time spent to learn and use the ASCLO-S editor have been also obtained. Although the subjects have not manifested any particular problems to learn the tool an issue has been arisen from Subject3, Subject5, Subject6, and Subject7. In particular, the issue concerns the binding between symbolic names and learners' stereotypes. Hence, they have suggested of improving the tool support to simplify this binding.

5.2. *Empirical analysis results*

Table 6 reports the descriptive statistics for the variables we have used to collect information on the e-learning processes. In particular, the first column

Author's personal copy

shows the Obs, while the remaining columns present for each variable the *MIN*, the *MAX*, the *MEAN*, and the standard deviation values (*STD. DEV.*) of these observations.

In order to demonstrate that the technical competences of an instructional designer do not influence the effort to use ASCLO-S editor, we have first analysed the correlation between the effort to define adaptive e-learning processes and the competences of the subjects and then the correlation between the development effort and the variables of the adaptive e-learning processes.

Table 6
Descriptive statistics of *EFH* and variables

	Obs	MIN	MAX	MEAN	STD. DEV.
EFH	10	49	250	93.7	58.18
RC	10	8	152	46.4	42.50
Me	10	1	131	33.1	42.37
ALP	10	1	7	3.3	2.31
KF	10	4	18	8.9	4.46
Qs	10	3	18	8.2	5.85
St	10	17	256	77	98.00

Table 7
Correlations between the effort and subject competences

	Pearson correlations		Spearman's ρ	
	Statistic	Sign	Statistic	Sign
EFH-q11	0.384	0.274	0.106	0.770
EFH-q12	0.405	0.246	0.183	0.613

Concerning the correlation between the effort to define adaptive e-learning processes and the competences of the recruited subjects we can observe that they are not statistically correlated using both Spearman's ρ [50] and Pearson correlations [51]. In particular, the effort for defining e-learning processes is not influenced by the subjects' technical competences as the answers of the questions q11 and q12 are not correlated with the effort (as indicated by the *p*-values in Table 7). The scatter plot in Fig. 6A graphically describes the relationships between *EFH* and q11, while the relationship between *EFH* and q12 is shown in Fig. 6B.

In order to identify the variables best correlated with the development effort we have again used both Spearman's ρ [50] and Pearson correlations [51]. The achieved results state that *RC* is the only variable statistically correlated to the effort variable (*EFH*) at level 0.05 (see Table 8). On the other hand, to show that *RC* is a useful effort indicator we have applied an OLS regression analysis [19] by considering *EFH* as dependent variable and *RC* as independent variable. This technique allows us to determine the equation of a line interpolating data. This equation is used to predict the effort in terms of the number of required person-hours.

Beside the *linearity*, in order to apply OLS regression analysis the residuals *homoscedasticity* and *normality* assumptions have to be verified [19].

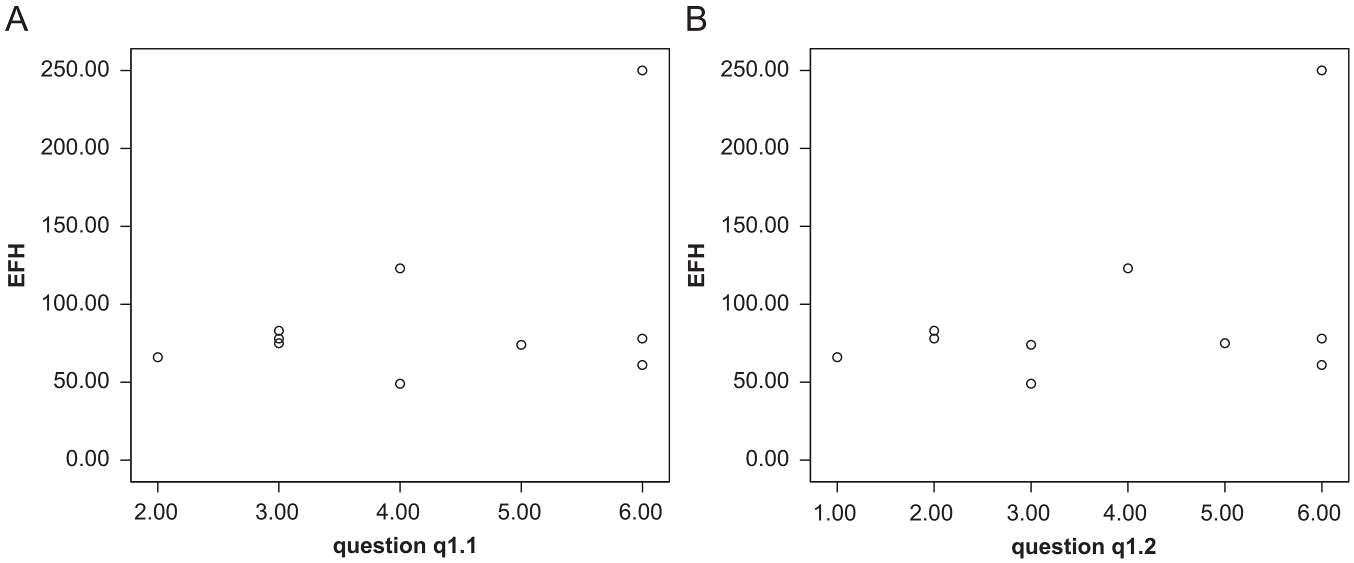


Fig. 6. The scatter plot of *EFH* vs. q11 (A) and *EFH* vs. q12 (B).

Author's personal copy

Table 8
Correlations between the effort and selected size measures

	Pearson correlations		Spearman's ρ	
	Statistic	Sign	Statistic	Sign
EFH-RC	0.957	0.000	0.796	0.006
EFH-ME	0.420	0.226	0.402	0.249
EFH-ALP	0.283	0.428	0.404	0.247
EFH-KF	0.196	0.586	0.628	0.052
EFH-Qs	0.220	0.541	0.214	0.553
EFH-St	0.220	0.541	0.428	0.217

Table 9
The Breush–Pagan homoscedasticity and the Shapiro–Wilk tests

Shapiro–Wilk normality test		Breush–Pagan homoscedasticity test	
Statistic	Sign	Statistic	Sign
0.891	0.173	0.713	0.398

The former establishes that the variance of the residuals has to be constant, while residuals are normally distributed in the latter assumption. To assess homoscedasticity and normality assumptions we have used the Breush–Pagan test [53] and the Shapiro–Wilk test [54], respectively. In particular, the null hypothesis of the Shapiro–Wilk test is that the sample is taken from a normal distribution, while the null hypothesis of Breush–Pagan test is that the variance of the residuals is constant. Thus, if the p-value is smaller than the critical value (usually 0.05) then the null hypothesis can be rejected. Concerning the Shapiro–Wilk test, we can assert that the population is not normal, while the residuals do not have constant variance for the Breush–Pagan test.

In our case, the Breush–Pagan test (see Table 9) suggests that the null hypothesis of homoscedasticity of the residuals cannot be rejected since the p-value of the statistic (i.e. *Sign*) is greater than 0.05, while the Shapiro–Wilk test indicates that the null hypothesis of normality of residuals cannot be rejected since the p-value of the statistic (i.e. *Sign*) is greater than 0.05.

The prediction model obtained by applying the linear regression analysis is

$$EFH = 1.311 \times RC + 32.881. \tag{M.1}$$

Table 10 shows the crucial indicators we have used to evaluate the quality of the resulting prediction model. It is worth noting that the model M.1 is characterised by a high R^2. The high F value and the low *Sign F* value suggest that the prediction is indeed possible with a high degree of confidence. Furthermore, the *RC* coefficient and the intercept are characterised by t-values and p-values satisfying the acceptable threshold values.

Table 10
Quality indicators of the model M.1

	t-value	p-value	R^2	F	*Sign F*
RC	9.382	0.000	0.917	83.03	0.000
Intercept	3.829	0.005			

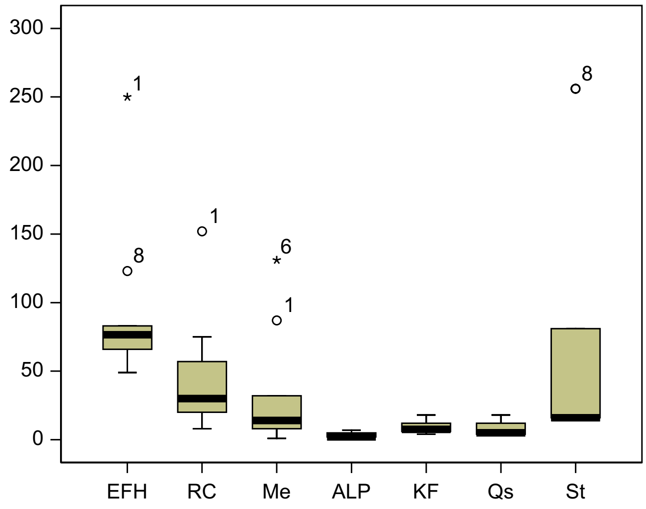


Fig. 7. The boxplots for the variables of Table 3.

In order to remove possible extreme values, which may improperly influence the identified model, the analysis of raw data has been performed. The analysis of *EFH* indicates that the observations 1 and 8 are possible outliers since they had values higher than the other ones (see Fig. 7). The analysis of the corresponding adaptive e-learning processes has revealed that it is reasonable justified by their size that is larger than the other ones. Thus, we have not removed these observations from the data set.

The analysis of the *RC* variable suggests that the observation 1 is a possible outlier. This is also an influential observation since it is characterised by a Cook's distance, which is greater than $4/10 = 0.4$. In this case, the observation that might influence the model should be removed and the model stability should be analysed by observing the effect of this removal. If the model coefficients remain stable and the R^2 improves the highly influential project is

Author's personal copy

retained in the data analysis. Therefore, we have removed the observation 1 and then we have analysed its effect on the obtained model, which follows:

$$EFH = 0.768 \times RC + 49.710. \tag{M.2}$$

The coefficients and the intercepts are significant in both the models M.1 and M.2, as shown in Tables 10 and 11. However, we can note that the values of the coefficients and the intercepts are substantially different. The values of R^2 and F of the two models are not closed. Moreover, the observation 1 improves the model fit as suggested by the R^2 value of the model M.1. As a consequence of the stability analysis, we have decided to carry out the empirical evaluation not considering the observation 1.

To assess the accuracy of the prediction model M.2 we have considered two evaluation criteria: *MMRE* and *Pred* (*l*) [52]. Table 12 shows both *MMRE* and *Pred*(0.25) values for the models obtained with linear regression analysis and the binary trees obtained by the regression tree technique.

The model M.2 exhibits an *MMRE* value less than 0.25 and a *Pred*(0.25) value greater than 0.75. These values suggest that the linear regression model can be considered appropriate for the effort prediction. To better understand the prediction power of the model M.2 we have also analysed the boxplot of the residuals in Fig. 8. This boxplot provides a good indication of the distribution of the residuals and improves the comprehension of *MMRE* and *Pred*(0.25), as suggested in [55]. The median for M.2 is closed to 0 and the boxplot has not outliers, thus indicating that the model provides a good effort estimation.

Table 11
Quality indicators of the model M.2

	t-value	*p*-value	R^2	*F*	*Sign F*
RC	3.910	0.006	0.686	15.29	0.006
Intercept	6.267	0.000			

Table 12
Accuracy evaluation

	MMRE	*Pred* (0.25)
M.2	0.16	0.78

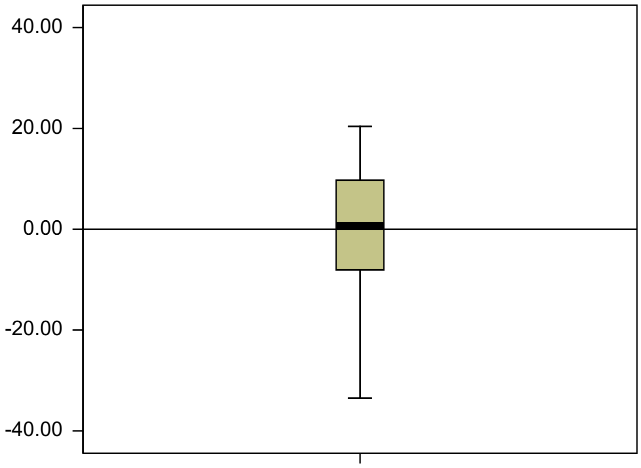


Fig. 8. The boxplots of residuals.

Concluding, the results of the empirical analysis have revealed that the effort required to model e-learning processes depends on the *RC* variable, but it is not influenced by the experience of the instructional designer.

6. Conclusions and discussion

Visual and diagrammatic representations play a central role in several application domains, since they are recognised to be important tools for describing and reasoning. Their employment allows us to improve productivity of expert and non-expert users in several application domains. It is because they provide means to easily capture and model difficult concepts. The "ease of use" of tools implementing visual languages is recognised to be one of the key aspects to improve the user satisfaction [11,56] that can be assessed performing usability studies [13,14].

The current trend of academic and industrial realities is to increase the use of e-learning. In particular, software tools supporting the critical task of instructional design should provide automated support for the analysis, design, documentation, implementation, and deployment of instruction via the Web. Moreover, instructional design should also aim at encouraging high self-expectations, satisfaction, and welfare by respecting the diversities among learners. To address these issues, visual language-based tools [5,6,57] as well as adaptive e-learning processes [2,58] and systems [26] could be adopted.

In this paper, we have presented and discussed a usability study of a visual language-based authoring

Author's personal copy

tool (i.e. ASCLO-S editor) conceived to define adaptive e-learning processes. The study is based on the combined use of two techniques: a questionnaire-based survey and an empirical analysis. The survey has revealed a good satisfaction degree of the subjects, who have not manifested any kind of listlessness and indifference during the use of the analysed tool. Encouraging results in terms of time spent to learn and use the tool have also been achieved. Concerning the empirical analysis, we have identified the answers that are statistically correlated to the subjects' technical competences using Spearman's ρ [50] and Pearson [51] correlations. We have also performed linear regression analysis [19,20] to identify effort estimation models based on some measures obtained from the definition of adaptive e-learning processes through the described visual language-based tool. The identified model has predicted with a good accuracy the effort to develop a learning process using *RC* variable. This effort indicator is considered relevant as the empirical analysis revealed. We have also showed that the effort required to define an e-learning process is not statistically correlated to the subjects' technical competences. As a consequence, the effort required to model e-learning processes depends on the size of the developed process but it is not influenced by the experience of the instructional designer with the use of e-learning tools. Concluding, the outcomes of the usability study showed that both the proposed visual notation and the system prototype are suitable for instructional designers with or without experience on tools for defining e-learning processes.

The usability study has also provided a number of directions to improve the ASCLO-S editor. A first direction would be to add some features to further support instructional designers in the binding between symbolic names and stereotypes. A second direction should aim at fixing some usability problems that the subjects properly identified.

Some directions for future work can be considered. The first one is suggested by the limitations of the presented study. Indeed, although OLS regression analysis requires from 5 to 10 observations for each variable in the model (see e.g. [59,60]), the use of a small group of subjects (nine in our case) may threat the validity of the empirical study. To this aim, we are planning to replicate the usability study using a larger number of subjects with different teaching experiences.

Future work will be also devoted to carry out an empirical study to compare the ASCLO-S editor with widely used commercial authoring tools, such as Macromedia Authorware. This study will also aim at corroborating the results of the empirical study presented in [18], where a preliminary comparison between ASCLO-S editor and Macromedia Authorware revealed that the former provides a better support for not expert users. It was also observed that Authorware is more suited for expert users and a specific background and computer science knowledge is required to master it. Finally, we also plan to use effort estimation models in the usability assessment of other e-learning tools.

References

[1] C. Ardito, M. De Marsico, R. Lanzilotti, S. Levialdi, T. Roselli, V. Rossano, M. Tersigli, Usability of E-learning tools, in: Proceedings of the Working Conference on Advanced Visual Interfaces, 2004, Gallipoli, Italy, pp. 80–84.

[2] G. Costagliola, A. De Lucia, F. Ferrucci, G. Scanniello, Adaptive learning processes in pure and blended teaching approaches, in: Proceedings of the IADIS International Conference: e-Society 2005, Qawra, Malta, 27–30 June 2005, IADIS Press, pp. 414–418.

[3] J. Cynthia, M. Finelli, A. Wicks, An instrument for assessing the effectiveness of the circuits curriculum in an electrical engineering program, IEEE Transaction on Education 43 (2) (2000) 137–142.

[4] C. Vrasidas, A systematic approach for designing hypermedia environments for teaching and learning, International Journal of Instructional Media (2002), 〈http://www.cait.org/vrasidas/hypermedia.pdf〉.

[5] M.A. Bell, D. Jackson, Visual author languages for computer-aided learning, in: Proceedings of the IEEE Workshop on Visual Languages, USA, 1992, pp. 258–260.

[6] J.D. Campbell, D.E. Mahling, A visual language system for developing and presenting Internet-based education, in: Proceedings of the IEEE Symposium on Visual Languages, Nova Scotia, Canada, 1998, pp. 66–67.

[7] A. Koohang, Expanding the concept of usability, in: Eli Cohen (Ed.), Informing Science: The International Journal of an Emerging Transdiscipline, vol. 7, 2004, pp. 129–141.

[8] A.F. Blackwell, Metacognitive theories of visual programming: what do we think we are doing?, in: Proceedings of the IEEE Symposium on Visual Languages, IEEE Computer Society Press, Los Alamitos, CA, 1996, pp. 240–246.

[9] K.N. Whitleym, A.F. Blackwell, Visual programming: the outlook from academia and industry, in: S. Wiedenbeck, J. Scholtz (Eds.), Proceedings of the Seventh Workshop on Empirical Studies of Programmers, 1997, pp. 180–208.

[10] A.F. Blackwell, T.R.G. Green, Does metaphor increase visual language usability?, in: Proceedings of the IEEE Symposium on Visual Languages, 1999, pp. 246–253.

[11] A. Dix, J. Finlay, G. Abowd, R. Beale, Human–Computer Interaction, second ed., Prentice-Hall, Englewood Cliffs, NJ, 1994.

Author's personal copy

[12] M. Matera, M.F. Costabile, F. Garzotto, P. Paolini, SUE inspection: an effective method for systematic usability evaluation of hypermedia, IEEE Transactions on Systems, Man and Cybernetics—Part A 32 (1) (2002) 93–103.

[13] J. Nielsen, Usability Engineering, Academic Press, New York, 1993.

[14] J. Nielsen, Usability testing, in: G. Salvendy (Ed.), Handbook of Human Factors and Ergonomics, Wiley, New York, 1997, pp. 1543–1568.

[15] K. Ricks, B.A. Arnoldy, How to conduct your own usability study, in: Proceedings of the International Professional Communication Conference, IEEE Computer Society Press, Portland, OR, USA, 2002, pp. 115–126.

[16] B. Shneiderman, Designing the User Interface, third ed., Addison-Wesley, Longman, Inc., Reading, MA, New York, 1998.

[17] G. Casella, G. Costagliola, F. Ferrucci, G. Polese, G. Scanniello, Visual languages for defining adaptive and collaborative e-learning activities, in: Proceedings of the IADIS International Conference: e-Society 2004, vol. 1, Avila, Spain, 16–19 July 2004, pp. 243–250.

[18] G. Costagliola, A. De Lucia, F. Ferrucci, G. Scanniello, Visual languages for non expert instructional designers: a usability study, in: Proceedings of the 11th International Conference on Distributed Multimedia Systems, Banff, Canada, September 5–7, Knowledge Systems Institute Press, 2005, pp. 46–51.

[19] D. Montgomery, E. Peck, G. Vining, Introduction to Linear Regression Analysis, third ed., Wiley, New York, 2001.

[20] L. Briand, K.E. Emam, D. Surmann, I. Wiekzorek, K. Maxwell, An assessment and comparison of common software cost estimation modeling techniques, in: Proceedings of the International Conference on Software Engineering, Los Angeles, USA, 16–22 May 1999, pp. 313–322.

[21] L. Briand, T. Langley, I. Wiekzorek, A replicated assessment and comparison of common software cost modeling techniques, International Software Engineering Research Network Technical Report ISERN-99-15.

[22] E. Mendes, I. Watson, C. Triggs, N. Mosley, S. Counsell, A comparative study of cost estimation models for web hypermedia applications, Empirical Software Engineering 8 (2) (2003) 163–196.

[23] D.J. Muraida, J.M. Spector, Automatic design instruction, in: S. Dijkstra, N. Seel, F. Schott, D. Tennyson (Eds.), Instructional Design: International Perspectives, vol. 2, Lawrence Erlbaum, Mahwah, NJ, 1997.

[24] P. Goodyear, Instructional design environments: methods and tools for the design of complex instructional systems, in: S. Dijkstra, N. Seel, F. Schott, R.D. Tennyson (Eds.), Instructional Design: International Perspectives, vol. 2, Lawrence Erlbaum Associates, Mahwah, NJ, 1997.

[25] P. Goodyear, Seeing learning as work: implications for understanding and improving analysis and design, Journal of Courseware Engineering 2 (1999) 3–11.

[26] AIMS: Adaptive Information System for Management of Learning Content, 2004, from 〈wwwis.win.tue.nl:8080/AIMS〉.

[27] E. Cloete, P. Kotzé, Designing interactive instructional authoring tools, Technical Report TR-UNISA-2003-01, 2003, available at 〈http://www.cs.unisa.ac.za/TechnicalReports/TR-UNISA-2003-01.PDF〉.

[28] C. Ardito, M.F. Costabile, M. De Marsico, R. Lanzilotti, S. Levialdi, P. Plantamura, T. Roselli, V. Rossano, M. Tersigli, Towards guidelines for usability of e-learning applications, in: Proceedings of the User Interfaces for All, Austria, Vienna, 2004, pp. 185–202.

[29] M.F. Costabile, M. De Marsico, R. Lanzilotti, V.L. Plantamura, T. Roselli, On the usability evaluation of e-learning applications, in: Proceedings of the 38th Annual Hawaii International Conference on System Sciences (HICSS'05), Waikoloa, Hawaii, USA, January 3–6, 2005.

[30] O. Parlangeli, E. Marchigiani, S. Bagnara, Multimedia System in Distance Education: Effects on Usability, Interacting with Computers, vol. 12, Elsevier, Great Britain, 1999, pp. 37–49.

[31] P. Zaharias, A usability evaluation method for e-learning: focus on motivation to learn, in: Proceedings of the Conference on Human Factors, CHI '06 Extended Abstracts on Human Factors in Computing Systems, Montréal, Québec, Canada, April 22–27, 2006, pp. 1571–1576.

[32] A.J. Albrecht, Measuring application development productivity, in: SHARE/GUIDE: Proceedings of the IBA4 Applications Development Symposium, 1979, pp. 83–92.

[33] A.J. Albrecht, J.E. Gaffhey, Software function, source lines of code, and development effort prediction, IEEE Transactions on Software Engineering 9 (6) (1983) 639–648.

[34] B. Boehm, Software Engineering Economics, Prentice-Hall, Englewood Cliffs, NJ, 1981.

[35] L.H. Putnam, A general empirical solution to the macro software sizing and estimation problem, IEEE Transactions on Software Engineering 14 (4) (1978) 345–381.

[36] G.R. Finnie, G.E. Wittig, A comparison of software effort estimation techniques: using function points with neural networks, case based reasoning and regression models, Journal of Systems and Software 39 (1997) 281–289.

[37] M. Jorgensen, Experience with the accuracy of software maintenance task effort prediction models, IEEE Transactions on Software Engineering 21 (8) (1995) 674–681.

[38] M. Shepperd, C. Schofield, Estimating software project effort using analogies, IEEE Transactions on Software Engineering 23 (12) (1997) 736–743.

[39] E. Stensrud, I. Myrtveit, Human performance estimation with analogy and regression models, in: Proceedings of the METRICS 98 Symposium, Bethesda, MD, USA, 1998, pp. 205–213.

[40] F. Walkerden, R. Jeffery, An empirical study of analogy-based software effort estimation, Empirical Software Engineering 4 (2) (1999) 135–158.

[41] L. Briand, I. Wieczorek, Software resource estimation, in: J.J. Marciniak (Ed.), Encyclopedia of Software Engineering, vol. 2 P-Z, second ed., Wiley, New York, 2002, pp. 1160–1196.

[42] K.M. Kapp, How Long Does It Take? Estimation Methods for Developing E-Learning, available at 〈http://www.learningcircuits.org/2003/jul2003/kapp.htm〉.

[43] R.P. Bodnar, S.M. Henry, Metrics for multimedia languages, in: Proceeding of the Fourth Conference on Software Maintenance and Reengineering, Zurich, Switzerland, February 29–March 3, 2000, pp. 75–82.

[44] G. Casella, G. Costagliola, F. Ferrucci, G. Polese, G. Scanniello, A SCORM thin client architecture for e-learning systems based on web services, International Journal of Distance Education Technologies, to appear.

Author's personal copy

[45] Advanced Distributed Learning (ADL), Sharable Content Object Reference Model (SCORM), 2006, 〈http://www.adlnet.gov/scorm/index.cfm〉.

[46] Advanced Distributed Learning (ADL) Initiative, 2006 〈http://www.adlnet.org〉.

[47] A.N. Oppenheim, Questionnaire Design, Interviewing and Attitude Measurement, Pinter Publishers, 1992.

[48] A.J.C. Cowderoy, Measures of size and complexity for web-site content, in: Proceedings of the Combined 11th European Software Control and Metrics Conference, Munich, Germany, 2000, pp. 423–431.

[49] B.A. Kitchenham, L.M. Pickard, S.J. Linkman, An investigation of the analysis techniques for software datasets, in: Proceedings of the Sixth International Symposium on Software Metrics, IEEE Computer Society Press, Los Alamitos, CA, 1999.

[50] C. Spearman, The proof and measurement of association between two things, American Journal of Psychology 15 (1904) 72–101.

[51] J.E. Freund, in: Mathematical Statistics, fifth ed., Prentice-Hall, Upper Saddle River, NJ, 1992, pp. 494–546.

[52] D. Conte, H.E. Dunsmore, V.Y. Shen, Software Engineering Metrics and Models, The Benjamin/Cummings Publishing Company, Inc., Menlo Park, CA, 1986.

[53] T.S. Breush, A.R. Pagan, A simple test for heteroscedasticity and random coefficient variation, Econometrica 47 (1979) 1287–1294.

[54] P. Royston, An extension of Shapiro and Wilk's test for normality to large samples, Applied Statistics 31 (2) (1982) 115–124.

[55] B.A. Kitchenham, L.M. Pickard, S.G. MacDonell, M.J. Shepperd, What accuracy statistics really measure, IEE Proceedings—Software 148 (3) (2001) 81–85.

[56] International Organization for Standardization, Ergonomic requirements for office work with visual display terminals. Part 11: Guidance on Usability, 1998.

[57] Macromedia, Authorware 7.0, 2005, from 〈http://www.macromedia.com/software/authorware/〉.

[58] Advanced Distributed Learning (ADL), SCORM Sequencing and Navigation (SN) Book, 2004, from 〈http://www.adlnet.org/screens/shares/dsp_displayfile.cfm?fileid=998〉.

[59] J. Kliijnen, Sensitivity analysis and related analyses: a survey of statistical techniques, Journal of Statistical Computation and Simulation 57 (1–4) (1997) 111–142.

[60] T. Menzies, C. Zhihao, J. Hihn, K. Lum, Selecting best practices for effort estimation, IEEE Transactions on Software Engineering 32 (11) (2006) 883–895.