International Journal of Human-Computer Interaction
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hihc20

Developing a Usability Evaluation Method for e-Learning Applications: Beyond Functional Usability

Panagiotis Zaharias [a] [b] & Angeliki Poylymenakou [b]

[a] Department of Product and Systems Design Engineering , University of the Aegean , Syros, Hermoupolis, Greece

[b] Department of Management Science and Technology , Athens University of Economics and Business , Athens, Greece

Published online: 14 Jan 2009.

INTL. JOURNAL OF HUMAN–COMPUTER INTERACTION, *25*(1), 75–98, 2009

ISSN: 1044-7318 print / 1532-7590 online
DOI: 10.1080/10447310802546716

Developing a Usability Evaluation Method for e-Learning Applications: Beyond Functional Usability

Panagiotis Zaharias[1,2] and Angeliki Poylymenakou[2]
[1]Department of Product and Systems Design Engineering, University of the Aegean, Syros, Hermoupolis, Greece
[2]Department of Management Science and Technology, Athens University of Economics and Business, Athens, Greece

In this article, the development of a questionnaire-based usability evaluation method for e-learning applications is described. The method extends the current practice by focusing not only on cognitive but also affective considerations that may influence e-learning usability. The method was developed according to an established methodology in HCI research and relied upon a conceptual framework that combines Web and instructional design parameters and associates them with the most prominent affective learning dimension, which is intrinsic motivation to learn. The latter is proposed as a new usability measure that is considered more appropriate to evaluate e-learning designs. Two large empirical studies were conducted in order to evaluate usability of e-learning courses offered in corporate environments. The results provide valuable evidence for reliability and validity of the method, thus providing evidence that usability practitioners can use it with confidence when evaluating the design of e-learning applications.

1. THE NEED TO DEVELOP A USABILITY EVALUATION METHOD FOR E-LEARNING APPLICATIONS

Organizations and educational institutions have been investing in information technologies to improve education and training at an increasing rate during the last two decades. Particularly in corporate settings, continuous education and training for the human resource is critical to an organization's success. Electronic learning (e-learning) has been identified as an enabler for people and organizations

The author wishes to thank all the trainees–participants in the empirical studies, as well as all the partners in the international project during pilot trial 1 and the Greek private bank (pilot trial 2) for their collaboration.

Correspondence should be addressed to Panagiotis Zaharias, Department of Management Science and Technology, Athens University of Economics and Business, 76, Patission str., Athens, 104 34, Greece. E-mail: pz@aueb.gr

to keep up with changes in the global economy. Within the context of corporate training, e-learning refers to training delivered on a computer that is designed to support individual learning or organizational performance goals (Clark & Mayer, 2003). Although e-learning is emerging as one of the fastest organizational uses of the Internet (Harun, 2002), most e-learning programs exhibit higher dropout rates when compared with traditional instructor-led courses. Evidence from recent literature from the corporate training arena (Bonk, 2002; Moshinskie, 2001) reveals high attrition rates (as high as 80% as reported in Hodges, 2004) in e-learning courses. A major contributing factor is that the learners are not motivated or cannot sustain their motivation in the courses. There are many reasons that can explain the high dropout rates, such as relevancy of content, comfort level with technology, availability of technical support, etc., but one major contributor is the poor design and usability of e-learning applications. The latter is the focal point of this study.

Evaluating the usability of e-learning applications is not a trivial task. An increase in the diversity of learners, technological advancements, and radical changes in learning tasks (learner interaction with a learning/training environment is often a one-time event) present significant challenges and render the possibility of defining the context of use of e-learning applications. Identifying whom the users are and what the tasks are in an e-learning context impose extra difficulties. In the case of e-learning design the main task for the user is to learn, which is rather tacit and abstract in nature (Zaharias & Poulymenakou, 2006). Actually there are no tasks in the typical sense; e-learning interaction goes beyond the task and work-related usability paradigm. As Notess (2001) argues, "Evaluating e-learning may move usability practitioners outside their comfort zone." Squires (1999) highlights the need for integration of usability and learning and points out the non-collaboration of workers in human–computer interaction (HCI) and educational computing areas. In fact, usability of e-learning designs is directly related to their pedagogical value. An e-learning application may be usable but not in the pedagogical sense, and vice versa (Albion, 1999; Quinn, 1996; Squires & Preece, 1999). Accordingly, usability practitioners need to familiarize themselves with the educational testing research, learning styles, and the rudiments of learning theory. It seems that there is an ellipsis of research-validated usability evaluation methods that address the user as a learner in a holistic way, which includes the consideration of cognitive and affective learning factors.

The purpose of this research is to develop and empirically test a questionnaire-based usability evaluation method for e-learning applications. This method is based on the most important affective learning factor—i.e., intrinsic motivation to learn—and addresses the user as a learner in a holistic way. It goes beyond the functional usability paradigm and proposes motivation to learn as a new type of e-learning usability measurement. The development of this method was based upon a very well-known methodology in HCI research and practice. Two large empirical studies were conducted examining usability of e-learning applications in authentic environments, where learners are employees in corporate settings. The focus was on empirical assessment of the reliability and validity of the method. The results provide evidence for reliability and validity of the method. A synopsis of the use of questionnaires in e-learning usability is presented in the next section, followed by the main body of this work, the description of method's development. The article concludes with a summary of results and limitations and discusses future research work.

1.1. Beyond Functional Usability: The Emergence of an Affective Dimension

This work aims at investigating e-learning design and evaluation dimensions beyond functional usability. In accordance with the user experience research agenda (Hassenzahl & Tractinsky, 2006), there is a strong need for empirical works that go beyond the purely cognitive and task-oriented paradigm. Affective and emotional aspects of interaction are key issues to explore while trying to understand and interpret the user experience (Gaver & Martin, 2000; Hassenzahl, 2003). This is in line with the contemporary affective HCI research where a major challenge is to address user affect. It is acknowledged that systems designers assess the range of possible affective states that users may experience while interacting with the system (Hudlicka, 2003), while at the same time new usability techniques and measures need to be established (Hornbaek, 2005). For instance, traditional usability measures of effectiveness, efficiency, and satisfaction are not adequate for new contexts of use such as home technology (Monk, 2002), ubiquitous computing (Mankoff et al., 2003), and e-learning (Soloway, Guzdial, & Hay, 1994; Zaharias, 2004). In the e-learning landscape, the traditional task and work-related usability seem to have limited value (Zaharias, 2004) while at the same time the need to approach the learner experience in a more holistic way becomes stronger. This challenge requires a focus on the affective aspects of learning (O'Regan, 2003; R. Picard, Kort, & Reily, 2001). It has been argued that affect is the "fuel" that learners bring to the learning environment connecting them to the "why" of learning. New developments in learning theories such as constructivist approaches put an increasing emphasis on the affective domain of learning; new thinking in adult learning theory and practice stresses the need to enhance learners' internal priorities and drives that can be best described by motivation to learn. The latter, a concept intimately linked with learning (Schunk, 2000), is the most prominent affective learning factor that can greatly influence users' interactions with an e-learning application. Thus, intrinsic motivation to learn is proposed as the anchor for the development of a new usability measure for e-learning design. Consequently, the focus of this study responds to the need for affective considerations that shed light into the user (learner) experience:

- Importance of affect is critical and must be addressed in the e-learning context.
- Motivation to learn is selected in this study among other affective states and emotions.
- This study combines Web usability attributes with instructional design attributes that can explain and predict to a certain extent the learners' intrinsic motivation to learn.

2. RELATED WORK: THE USE OF QUESTIONNAIRES IN USABILITY STUDIES

Questionnaires have been widely used for evaluating usability of interactive systems. The main advantage is that usability questionnaires provide feedback from the point of view of the user; in addition, they are usually quick and cost effective to administer and to score. A large amount of data can be collected and such data can be used as a reliable source to check whether quantitative usability targets have been met. In the same vein, Root and Draper (1983) argued that questionnaire methodology clearly meets the requirements of inexpensiveness and ease of application. Such requirements are of great importance in e-learning context; e-learning practitioners need a short, inexpensive, and easy-to-deploy usability evaluation method so that e-learning economics can afford its use (Feldstein, 2002; Zaharias, 2004). There have been many efforts to develop usability questionnaires for software product evaluation such as the Questionnaire for User Satisfaction (QUIS), the Computer Satisfaction Inventory (CUSI), the Post-Study System Usability Questionnaire (PSSUQ), the Software Usability Measurement Inventory (SUMI), etc. However, these questionnaires are too generic (Keinonen, 1998; Konradt, Wandke, Balazs, & Christophersen, 2003), and the need for more specific questionnaires tailored to particular technologies or groups of software products has increased. In order to meet this crucial requirement, tailored questionnaires have been developed, such as Website Analysis and MeasureMent Inventory (WAMMI; Kirakowski, Claridge, & Whitehead, 1998) for Web site usability, Measuring Usability of Multi-Media Systems (MUMMS) for evaluating multimedia products, and the Usability Questionnaire for Online Shops (UQOS; Konradt et al., 2003) for measuring usability in online merchandisers.

As far as the development of usability questionnaires tailored to e-learning, much of the work conducted is either anecdotal or information about empirical validation is missing. Thus, it is no surprise that most e-learning researchers and practitioners usually employ some of the well-known and validated satisfaction questionnaires. Such questionnaires or variations of them have been used in several e-learning studies: Parlangeli, Marchigianni, and Bagnara (1999) and Chang (2002) used a variation of QUIS (Chin, Diehl, & Norman, 1988; Shneiderman, 1987) and Avouris, Tselios, Fidas, and Papachristos (2001) used a variation of WAMMI (Kirakowski et al., 1998). In addition, Avouris (1999) customized the QUIS usability questionnaire for evaluation of educational material. The questionnaire contains 75 questions grouped in 10 separate categories. These categories are (a) general system performance, (b) software installation, (c) manuals and on-line help, (d) on-line tutorials, (e) multimedia quality, (f) information presentation, (g) navigation, (h) terminology and error messages, (i) learnability, and (j) overall system evaluation. Other studies report some attempts toward the development of new questionnaires customized for e-learning; Wade and Lyng (1999) developed and used a questionnaire with the following criteria: (a) naturalness, (b) user support, (c) consistency, (d) nonredundancy, and (e) flexibility. Tselios, Avouris, Dimitracopoulou, and Daskalaki (2001) used a questionnaire with 10 items, which have been adapted from well-known heuristics. De Villiers (2004) developed a questionnaire by adapting "learning with software heuristics" proposed by Squires and Preece (1999). Such research efforts produce some outcomes toward the direction of addressing the user as a learner but any other details about the development and psychometric properties of questionnaires such as reliability and validity are missing. In addition, most of these developments focus mainly on cognitive aspects that influence the interaction with e-learning applications while they neglect or partly cover the affective ones, whose importance is continuously increasing.

On the other hand, numerous studies have employed questionnaires as tools for assessing an array of affective states such as motivation (Keller & Keller, 1989), curiosity, interest, tiredness, boredom, etc. (Whitelock & Scanlon, 1996), or achievement motive, creativity, sensation seeking, etc. (Matsubara & Nagamachi, 1996), but most of them do not relate such users' affective states with any usability evaluation effort or measurement. As already mentioned, empirical studies that relate affective-emotional aspects of interaction with users' perceptions of usability are lacking (Hassenzahl & Tractinsky, 2006).

According to the above analysis, the main difference between this usability questionnaire and other usability questionnaires can be summarized as follows: (a) The proposed usability questionnaire is built originally upon the most learning affective aspect: intrinsic motivation to learn. This is in contrast with the majority of the existing standardized usability questionnaires that have been developed around task-related and cognitive aspects and do not incorporate or partly cover some general affective aspects (Hornbaek, 2006). (b) It is exclusively made to meet e-learning specificities covering the double user-learner persona, whereas most of the existing standardized questionnaires are quite generic (Nokelainen, 2006; Ryu & Smith-Jackson, 2006). (c) The proposed questionnaire is under a rigorous process of empirical psychometric testing, whereas in most cases relevant studies that employ questionnaires in an e-learning context do not report any information about empirical psychometric validation (Nokelainen, 2006).

3. METHOD DEVELOPMENT

The questionnaire design methodology suggested by Kirakowski and Corbett (1990) was followed. This methodology contains five stages: (a) forming the survey, (b) item sampling, (c) pilot trial, (d) production version, and (e) next version. This article reports the work conducted along the first three stages, where reliability and validity of the proposed method were tested and established.

3.1. Forming the Survey

Regarding the "forming the survey" stage, two objectives were considered: (a) the type and (b) the scope of the questionnaire. A psychometric-type of questionnaire was considered as the most suitable method since the major objective was to measure users' perception of e-learning usability and related affect (in this case motivation to learn). As far as the scope of the questionnaire, it has to be noted that this research focuses on usability evaluation of asynchronous e-learning applications as adult learners use them in corporate settings for training purposes. In addition, this method extends the current practice by focusing on motivation to learn as an

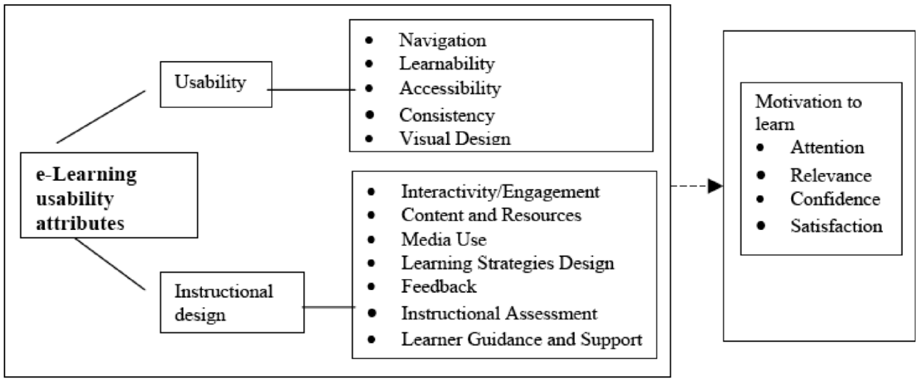

FIGURE 1 A theoretical framework for e-learning usability employing motivation to learn.

important affective learning dimension to address the user as a learner in a more holistic way.

This research relies mainly on a theoretical framework (Zaharias, 2004, 2005, 2008), which employs (a) a combination of Web design and instructional design parameters and (b) intrinsic motivation to learn (Keller, 1983), see Figure 1. Intrinsic motivation can be characterized as the drive arising within the self to carry out an activity whose reward is derived from the enjoyment of the activity itself (Csikszentmihalyi, 1975). The theoretical framework presupposes that all learning environments should foster intrinsic learning motivation (Spitzer, 1996). It further approaches the learner as a double persona (Smulders, 2002): the user of an e-learning application is a user that needs to interact with an interactive system and at the same time is a learner that needs to learn. The difference between users and learners can be boiled down to the issue of form (user interface) versus content (learning material). The user part of persona is concerned mostly with the form and the learner part is mostly interested in the content. This is not to say that content is not important in other contexts, such as e-commerce sites, e-shops, etc. Rather, this distinction aims at highlighting the special value that content brings to e-learning applications. Consequently, this framework incorporates Web usability parameters that satisfy the user part of the double persona and instructional design parameters that satisfy the learner part of the double persona and integrates them with intrinsic motivation to learn. According to the theoretical framework, the usability parameters were chosen from an array of studies for inclusion into the questionnaire such as: Navigation, Learnability, Accessibility, Consistency, Visual Design, Interactivity, Content & Resources, Media Use, Learning Strategies Design, Instructional Feedback, Instructional Assessment, and Learner Guidance & Support. All these parameters were carefully selected to associate them with learners' motivation to learn. Keller's (1983) theory and related model, which is perhaps the most influential model of motivation, was adopted in order to further analyze and interpret the motivation to learn construct. According to Keller (1983), the motivation to learn construct is composed of four subconstructs: attention, relevance, confidence, and satisfaction.

According to the above, a new set of usability parameters (derived from Web usability and instructional design literature) for e-learning is proposed that is directly associated with the most prominent affective learning factor: motivation to learn.

3.2. Item Sampling

The design parameters included in the theoretical framework were the main constructs included in the questionnaire. These constructs were measured with items adapted from prior research. To identify items for possible inclusion in the questionnaire, an extensive review of prior studies referring to e-learning and Web design was conducted. More specifically, a number of Web and instructional design guidelines (Johnson & Aragon, 2002; Lynch & Horton, 1999; Nielsen, 2000; Weston, Gandell, McApline, & Filkenstein, 1999) have been reviewed, as well as a number of usability evaluation heuristics, checklists, and questionnaires (Horton, 2000; Quinn, 1996; Reeves et al., 2002). Items were carefully selected to cover all the parameters included in the theoretical framework (see Appendix). The items in the questionnaire were presented in groups relating to each parameter; the aim of this questionnaire was to capture usability parameters that seem to have an effect on motivation to learn when measuring the usability of e-learning courses rather than to develop an equal scale of each parameter (i.e., parameters represented by an equal number of items). The questionnaire development started with an initial item pool of over 90 items. The items were examined for consistency of perceived meaning by getting three subject matter experts to allocate each item to content areas. Some items were eliminated when they produced inconsistent allocations. Prior to completion of the first version of the questionnaire, a mini pilot test was undertaken to ensure that items were adapted and included appropriately in the questionnaire. An online version of the questionnaire was administered to 15 users in academic settings (2 of them were usability experts) who had some prior experience with e-learning courses. Data were analyzed mainly for response completeness; some adjustments were made and subsequently some items were reworded. The whole procedure led to the first version of questionnaire, which consisted of 64 items, 54 items measuring Web and instructional design usability and 10 items measuring motivation to learn. Criteria corresponding to each usability parameter were assessed on a 5-point scale, where the anchors were 1 for *strongly disagree* and 5 for *strongly agree* (Table 1). Additionally, an option "Not Applicable" was included outside the scale for each item included in the questionnaire. Such an option has been used in several usability questionnaires (Lewis, 1995). There was also space for free-form comments. The first version of questionnaire was tested in pilot trial 1.

3.3. Pilot Trial 1

The first version of the questionnaire was used and empirically tested during an international research project. This project aimed at enhancing information and

Table 1: An Excerpt of the Questionnaire Version 3

Criteria	*1 Strongly disagree*	*2 Disagree*	*3 Neutral*	*4 Agree*	*5 Strongly agree*	*NA*
Content						
Vocabulary and terminology used are appropriate for the learners						
Abstract concepts (principles, formulas, rules, etc.) are illustrated with concrete, specific examples						
Learning & Support						
The courses offer tools (taking notes, job aids, recourses, glossary, etc.) that support learning						
The courses include activities that are both individual based and group based						
Visual Design						
Fonts (style, color, saturation) are easy to read in both on-screen and printed versions						
Navigation						
Learners always know where they are in the course						
The courses allow the learner to leave whenever desired but easily return to the closest logical point in the course						
Accessibility						
The course is free from technical problems (hyperlink errors, programming errors, etc.)						
Interactivity						
The courses use games, simulations, role-playing activities, and case studies to gain the attention and maintain motivation of learners						
Self-Assessment & Learnability						
Learners can start the course (locate it, install plug-ins, register, access starting page) using only online assistance						
Motivation to Learn						
The course incorporates novel characteristics						
The course stimulates further inquiry						
The course is enjoyable and interesting						
The course provides learner with frequent and varied learning activities that increase learning success						

communication technologies' skills (ICT) in Southeastern (SE) European countries. The strategic goal of this project was to set up an e-learning service that provides e-learning courses on ICT skills and competences. Four user organizations representing four different countries in the region participated in the project: Greece, Romania, Bulgaria, and Turkey. The Greek and Turkish organizations are operating in the IT services industry, the Bulgarian in the telecommunications industry, and the Romanian organization in the power and petroleum industry. These user organizations used the e-learning service in order to train a subset of their employees (mostly IT professionals) in the following topics: IT Business Consultancy and Software and Applications Development. The main pillars of the e-learning service developed during the project were two asynchronous e-learning courses covering the two aforementioned topics.

Subjects and method for testing. A summative type of evaluation was conducted after the implementation of the e-learning service. Usability evaluation of the e-learning courses was a main component of the summative evaluation. The trainees (employees of user organizations) that participated in the project interacted with the e-learning service and the e-learning courses during a period of 4 months and then they were asked to participate in the summative evaluation study. In this study, a Web-based version of the questionnaire was developed, which is particularly useful in Web usability evaluation when the users are geographically dispersed (Tselios et al., 2001). The online questionnaire targeted the whole trainee population who used the e-learning service and it was released right after the end of the project's formal training period. The questionnaire was uploaded on the corporate intranets of the four user organizations and the human resources department of each organization was responsible for controlling and validating the whole process. The respondents were asked to evaluate the e-learning courses they had already used and interacted with. They self-administered the questionnaire and for each question were asked to circle the response that best described their level of agreement with the statements. The total number of trainees who participated in usability evaluation was 113 (63 male and 50 female); 36.27% were involved in software engineering, 9.80% in systems administration, and 18.63% in other ICT-related activities. In total 64.18% of respondents were responsible for technical tasks and responsibilities, and 14,71% were involved in managerial tasks.

Analysis and results. As already mentioned, the purpose of this research was to develop and empirically test a questionnaire-based usability evaluation method for e-learning applications. While constructing a questionnaire, a major objective is to end up with a concise and not very long (in terms of number of items) or complicated instrument. Therefore, a factor analysis was conducted. The general purpose of factor analysis is to find a way to condense the information contained in a number of original variables (constructs or items) into a smaller set of composite dimensions with a minimum loss of information (Hair, Anderson, Tatham, & Black, 1998). In addition, factor analysis can identify the structure of relationships among variables (i.e., identify dimensions that are latent). Consequently, as a first step for the refinement of the instrument, a factor analysis was conducted in order to identify the underlying dimensions of usability parameters of e-learning courses as perceived by the trainees. The Kaiser-Mayer-Olkin (KMO) Measure of Sampling Adequacy (which indicates whether the sample size is adequate for performing factor analysis and varies from 0 to 1.0) was 0.846, which is comfortably higher than the recommended level of 0.6 (Hair et al., 1998). The following rules were applied during the factor analysis: (a) A principal components extraction (one of the methods for factors' extraction, which is generally used when the purpose is data reduction) with Varimax rotation, which is the most common rotation method (rotation serves to make the output more understandable and is usually necessary to facilitate the interpretation of factors);

Table 2: Factor and Reliability Analysis in Pilot Trial 1

Factors	*Reliability Cronbach's alpha*	*Eigenvalue*	*Percentage of variance explained*
Interactive Content	∞ = 0.901	15.229	36.259
Instructional Feedback & Assessment	∞ = 0.852	3.090	7.357
Navigation	∞ = 0.801	2.341	5.573
Visual Design	∞ = 0.793	1.798	4.280
Learner Guidance & Support	∞ = 0.805	1.549	3.688
Learning Strategies Design	∞ = 0.821	1.354	3.223
Accessibility	∞ = 0.718	1.169	2.784
Learnability	∞ = 0.707	1.134	2.700
Percentage of total variance explained			65.865

(b) a minimum eigenvalue (which represents the amount of variance accounted for by a factor) of one was used as a cutoff value for extraction; (c) items with factor loadings less than 0.32 on all factors or greater than 0.32 on two or more factors were deleted. An eight-factor solution was extracted explaining 65.865% of the variance (Table 2 and Appendix). The whole process led to the refinement of the questionnaire and a more parsimonious solution has been reached, with 8 factors representing usability parameters of e-learning courses as follows: Interactive Content, Instructional Feedback & Assessment, Navigation, Visual Design, Learner Guidance & Support, Learning Strategies Design, Accessibility, and Learnability. In addition, factor analyses led to a reduced set of variables (i.e., items in the questionnaire). The first version of the questionnaire contained 54 items representing 12 usability parameters and 10 items representing the motivation to learn construct, for a total of 64 items. The second version of the questionnaire contained 41 it ems representing 8 usability parameters (the 8 factors extracted as already presented) and 10 items representing the motivation to learn construct, for a total of 51 items.

The factor analysis served as the basis for the reliability analysis, which was the second analysis toward the instrument's refinement. Reliability analysis is an assessment of the degree of consistency between multiple measurements of a variable. The most commonly used measure of reliability is internal consistency (i.e., consistency among the variables in a summated scale). In order to assess the internal consistency of the factors scales, Cronbach's alpha (the most widely used measure) was utilized.

As Table 2 exhibits, all factors show high internal consistency as indicated by high alpha coefficients (ranges from 0.707 to 0.901), which exceed the recommended level of .70 (Hair et al., 1998; Lewis, 1995). In addition, the composite variable motivation to learn shows a very high internal consistency, as the alpha coefficient indicates ($\alpha = 0.926$).

The new version of the questionnaire, which shows a very high reliability as measured by Cronbach's alpha (0.961), was used and empirically tested for further refinement in the second large-scale study (pilot trial 2) to be described in the next section.

3.4. Pilot Trial 2

The second large-scale study was conducted in collaboration with a large private Greek organization that operates in the banking industry. The organization has an extensive network of branches all over Greece and trainees were geographically distributed. The organization had implemented e-learning services and numerous e-learning courses during the last 2 years. The total number of e-learning courses was 23. Some were offered only through the e-learning mode (asynchronous e-learning courses) and others were offered as supplementary to traditional, classroom-based courses.

Subjects and method for testing. The second version of the questionnaire was paper based and it was delivered in the Greek language. The target population of the study was the whole set of employees who were trained via the e-learning courses. Thus, each participant in the study had experience in using the e-learning courses. The questionnaire version 2 was sent to the respective branches and the respondents (employees of these branches) were asked to evaluate the e-learning courses they had already used and interacted with. They self-administered the questionnaire and for each question were asked to circle the response that best described their level of agreement with the statements. Out of the 500 questionnaires distributed, 260 complete responses were returned for a response rate of 52%. Four responses were considered as invalid while cleaning the data; therefore, data from 256 respondents were actually used in the analysis. Among them 110 were male and 146 were female.

Analysis and results. A second round of statistical analyses was conducted in order to further test and refine the instrument. These analyses were the same as those performed in pilot trial 1. First a factor analysis was performed: Forty-one items representing eight usability parameters were factor analyzed using principal components method with a Varimax rotation. The Kaiser-Mayer-Olkin (KMO) measure of sampling adequacy was 0.914, which is comfortably higher than the recommended level of 0.6 (Hair et al., 1998). The following criteria were used in extracting the factors: a factor with an eigenvalue greater than one and factors that account for a total variance at least 50% would be selected. Further, only items with a factor loading greater than 0.32 would be included in each factor grouping. After performing two iterations of factor analysis, the pattern of factor loadings suggested that a seven-factor solution should be extracted. The seven-factor solution explained 54% of the variance. These factors were Content, Learning & Support, Visual Design, Navigation, Accessibility, Interactivity, and Self-Assessment & Learnability. In order to assess the internal consistency of the factors' scales, Cronbach's alpha was utilized.

According to the above analysis, it is evident that all factors show high internal consistency (Table 3) as indicated by high alpha coefficients, which exceed the recommended level of 0.70 (Hair et al., 1998; Lewis, 1995). In addition, the composite variable motivation to learn shows a very high internal consistency, as alpha coefficient indicates ($\alpha = 0.900$). After conducting factor and reliability analysis, a

Table 3: Factor and Reliability Analysis in Pilot Trial 2

Factors	*Reliability Cronbach's alpha*	*Eigenvalue*	*Percentage of variance explained*
Content	α = 0.824	12.277	30.691
Learning & Support	α = 0.868	2.530	6.326
Visual Design	α = 0.775	1.684	4.211
Navigation	α = 0.762	1.510	3.774
Accessibility	α = 0.734	1.313	3.282
Interactivity	α = 0.782	1.194	2.986
Self-Assessment & Learnability	α = 0.724	1.087	2.716
Percentage of total variance explained			53.986

new, more parsimonious solution emerged and a new version (version 3) of the questionnaire was derived. The overall alpha for the new version of questionnaire was very high, α = 0.934. This version of the questionnaire contains 39 items measuring e-learning usability parameters plus 10 items measuring motivation to learn, for a total of 49 items (Table 1 presents an excerpt of the questionnaire version 3).

4. VALIDITY ANALYSIS

The validity of a questionnaire concerns what the questionnaire measures and how well it does so. There are three types of validity: content validity, criterion-based validity, and construct validity (Anastasi & Urbina, 1997). Unlike criterion validity, evidence for construct validity cannot be obtained from a single study but from a number of interrelated studies (Saw & Ng, 2001). In this study, efforts were made to assess content and criterion validity of the questionnaire.

Regarding content validity, the usability attributes were thoroughly examined and chosen based on the HCI and more specifically on Web usability literature and instructional design literature. An extensive literature review was conducted in order to select the appropriate usability attributes; items for inclusion within the questionnaire were selected from a wide range of Web course design guidelines, checklists, and questionnaires. The factor analyses that were performed on data collected during trial 1 and trial 2 support the content validity, since meaningful unitary constructs emerged.

Furthermore, criterion validity can take two forms: concurrent and predictive validity. Concurrent validity is a statistical measure in conception and describes the correlation of a new instrument (in this case the questionnaire) with existing instruments, which purport to measure the same construct (Rust & Golombok, 1989). No attempt to establish concurrent validity was done. The main reason is the lack of other research-validated usability questionnaires that measure motivation to learn for cross-validating purposes.

Regarding predictive validity, a multiple regression analysis was performed during pilot trial 1. The main objective was to assess the efficacy and effectiveness

Table 4: Usability Parameters Predicting Motivation to Learn

Model summary				
Model	*R*	*R^2*	*Adjusted R^2*	*Sig. F change*
1	.846	.716	.691	.000

Predictors: (Constant), Learnability, Instructional Feedback & Assessment, Navigation, Accessibility, Learning Strategies Design, Visual Design, Interactive Content, Learner Guidance & Support

of the proposed usability parameters in explaining and predicting motivation to learn. In this research, the proposed usability parameters (i.e., the factors identified in factor analysis) are the independent variables (IVs) and the composite variable motivation to learn is the dependent variable (DV). The composite dependent variable consisted of the 10 items used to measure motivation to learn. All independent variables were entered into the analysis simultaneously in order to assess the predictive strength of the proposed model. When all independent variables were entered into the multiple regression model, results showed an R^2 of 0.716 and adjusted R^2 of 0.691 (Table 4). An analysis of variance revealed an F (8,93) of 29.264, which is statistically significant ($p < .001$). These findings reveal that the eight usability parameters (factors extracted from factor analysis in pilot trial 1), when entered together in the regression model, accounted for 71.6% (adjusted R^2 69.1%) of the variance in motivation to learn. Such findings practically support the main argument of this work, which is that perception of e-learning usability (in terms of the main identified usability dimensions; i.e., the factors) is strongly related with learners' motivation to learn (or actually can predict learners' motivation to learn). Such good results can be considered as preliminary evidence of the validity of the proposed method.

Construct validity has two components (Anastasi & Urbina, 1997; Rust & Golombok, 1989): *convergent validity*, which demonstrates association with measures that are or should be related and *divergent* or *discriminant validity*, which demonstrates a lack of association with measures that should not be related. Unlike criterion validity, evidence for construct validity cannot be obtained from a single study but from a number of interrelated studies (Saw & Ng, 2001). Nevertheless, a first attempt to assess construct validity was made within the context of this research. A variation of Multitrait Multimethod Matrix technique as suggested by Campbell and Fiske (1959) was used. According to this technique, convergent validity tests whether the correlations between measures of the same factor are different than zero and large enough to warrant further investigation of discriminant validity. In order to assess convergent validity, the smallest within-factor correlations are estimated. The correlation analysis revealed the smallest within-factor correlations: Content = 0.230; Learning & Support = 0.346; Visual Design = 0.411; Navigation = 0.289; Accessibility = 0.486; Interactivity = 0.264; and Self-Assessment & Learnability = 0.326. These correlations are significantly different than zero ($p < 0.01$) and large enough to proceed with discriminant validity analysis.

Discriminant validity for each item is tested by counting the number of times that the item correlates higher with items of other factors than with items of its own theoretical factor. For discriminant validity, Campbell and Fiske (1959) suggest that the count should be less than one half the potential comparisons. According to the correlation matrix and the factor structure there are 690 potential comparisons. A careful examination of correlation matrix revealed 356 violations of the discriminant validity condition, which is slightly higher than the one half the potential comparisons (345). This cannot be considered as a major problem since—as already mentioned—construct validity cannot be obtained from a single study but from a number of interrelated studies. Therefore, further examination of construct and, more specifically, divergent validity is an issue of further research.

4.1. External Validity

Analysis of data derived from pilot trial 2 pointed toward a refined version of the questionnaire, version 3. Analysis and results revealed that several usability attributes (the seven factors as extracted from factor analysis in pilot trial 2) of the third version of the questionnaire exhibit quite high and acceptable reliability and the overall reliability of the questionnaire is very high ($\alpha = 0.921$). This means that the questionnaire has a high internal consistency. Regarding validity, it was shown that content validity, criterion validity, and convergent validity have been adequately achieved. A final but also important note must be made concerning the external validity of the proposed method. External validity is about generalizing and has long been accentuated in usability literature (John & Marks, 1997; Thomas & Kellogg, 1989; Wolf, Caroll, Landauer, John, & Whiteside, 1989). Usability practice has borrowed relevant concepts from experimental psychology where the concept of ecological validity is important. This refers to what extent it is possible to generalize from experimental settings to real-world situations. This is an important consideration while developing and using a usability evaluation method. What endangers the ecological validity of a usability method's results is the fact that the test situation removes both the surroundings and the dynamics of the activity in the real world.

Research design in this study was set up in order to avoid such problems. The actual testing of the questionnaire-based usability evaluation method was conducted in corporate e-learning conditions in both pilot trials. Participants in both pilot trials were employees who were trained by using e-learning courses for corporate training purposes in their work environments with the appropriate support. In addition, in both trials there was plenty of time for interaction with the e-learning courses and appropriate external motivational characteristics were in place (e.g., certificates of successful attendance and completion of e-learning courses). To this end it can be supported that external validity requirements have been met and the results of the questionnaire-based usability evaluation method can be generalized to other user populations and organizational settings.

All the above findings led to the fourth stage, "production version," of the questionnaire design methodology that was followed throughout this study: "At this stage, the questionnaire is a mature one, and can be used with safety in research, since its reliability and validity have been established in the pilot runs" (Kirakowski & Corbett, 1990, p. 214).

5. DISCUSSION AND FURTHER WORK

5.1. Summary of the Results and Limitations

Motivated by the need to address the specificities of e-learning design, a usability evaluation method was developed. The proposed questionnaire-based usability evaluation method extends the current functional usability measures for e-learning and focuses on a more holistic user (learner) experience. It gives emphasis on the positive users' affective engagement and proposes motivation to learn as a new type of e-learning usability measurement.

The following statistical analyses were performed in order to test the main psychometric properties (reliability and validity) of the proposed instrument:

1. First a factor analysis was conducted in order to identify the underlying dimensions (i.e., the factors) of e-learning usability and reduce the length of the instrument (the number of items contained in the questionnaire).
2. After factor analysis, reliability analysis was performed to test the internal consistency of the identified factors.

Such analyses were performed in both pilot trials. Results revealed that overall internal consistency of the questionnaire is very high, and content and criterion validity have been adequately achieved.

Additionally, a multiple regression analysis was performed in order to investigate the relationship between the construct of motivation to learn and the e-learning usability dimensions (as they were identified from factor analysis). Results verified a strong relationship between e-learning usability and motivation to learn. The latter is proposed as a new type of e-learning usability measurement and the questionnaire is proposed as a proper instrument toward this goal.

There are still certain limitations that practitioners should be aware of. The questionnaire was focused only on asynchronous e-learning applications since this kind of delivery mode is the dominant approach. Finally, further studies are needed to provide more empirical evidence regarding construct validity of the method. It is widely recognized that validation is a process (Hornbaek, 2005) and adequate evidence for construct validity cannot be obtained from a single study but from a number of interrelated studies (Anastasi & Urbina, 1997; Saw & Ng, 2001).

5.2. Future Studies

Firstly, future studies can be designed in order to address the above limitations. As already mentioned, using a questionnaire as a method to assess an affective state has advantages and some weaknesses as well. Future research efforts can employ a combination with other methods to gather other information about more transient characteristics and more qualitative usability data. A combination of methods can give stronger results. Such methods could include the use of expert systems and sentic modulation, which is about detecting affective states through sensors such as cameras, microphones, wearable devices, etc. (Picard & Daily, 2005).

Moreover, further consideration is needed to explore usability attributes and the role of affect in synchronous e-learning courses and environments. Synchronous e-learning is characterized by different types of interactions through chats, real-time audio, application sharing, whiteboards, videoconferencing, etc., which imposes additional considerations concerning usability and its evaluation.

Regarding the additional evidence for reliability and validity, a modification of the third version of the questionnaire can be conducted by replacing and rewording some of the few items in the questionnaire that did not discriminate well and further test the method with different kinds of e-learning courses, different types of learners, and different corporate environments. Alternative statistical analyses can also be used to shed light on reliability and validity issues; for example, confirmatory factor analysis (CFA) can be used to determine convergent and divergent (or discriminant) validity (Wang, 2003). The advantages of applying CFA compared to classical approaches to determine convergent and divergent validity are widely recognized (Anderson & Gerbing, 1997).

Beyond the current limitations, future research can focus on the following:

- Use of the proposed questionnaire as a formative evaluation method: This article described the use of the questionnaire-based usability evaluation method in two large-scale empirical studies as a summative evaluation method. The proposed usability evaluation method can also provide useful design guidelines during the iterative design process as a formative evaluation method. Currently, the proposed usability evaluation method can point toward specific usability problems with an e-learning application. A more systematic exploitation of using such a method for formative evaluation can be realized through the development of a database where the results of a number of different usability studies can be stored so that the knowledge obtained can be reused.
- Benchmarking: The proposed questionnaire-based usability evaluation method can also provide benchmark information like other research-validated questionnaires (for example, WAMMI) do. This means that the usability of an e-learning application could be tested against others. A standardized database could be developed that contains the usability profiles of existing e-learning applications and, thus, can facilitate designers in comparing the usability of one application with a series of other e-learning applications.
- Next version of the questionnaire: focusing on other affective/emotional states. Future research should seek a deeper understanding of the design issues that influence learners' affect and emotions. Emotions such as fear, anxiety, apprehension, enthusiasm, and excitement as well as pride and embarrassment (Ingleton & O'Regan, 1998; O'Regan, 2003) along with the flow experience (Csikszentmihalyi, 1975, 1990; Konradt & Sulz, 2001) can provide significant input to e-learning design and shed light in explaining learners' behaviors; such emotions and their assessment can also be taken into consideration in the next version of the questionnaire.

REFERENCES

Albion, P. R. (1999). Heuristic evaluation of educational multimedia: From theory to practice. In *Proceedings of the 16th Annual conference of the Australasian Society for Computers in Learning in Tertiary Education.* http://www.ascilite.org.au/conferences/brisbane99/papers/albion.pdf

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Anderson, J. C., & Gerbing, D. W. (1997). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin 103*, *3*, 411–423.

Avouris, N. (1999). *Educational software evaluation guidelines* (Tech. Rep.). Athens: Working Group on Educational Software Evaluation Procedure, Pedagogical Institute.

Avouris, N., Tselios, N., Fidas, C., & Papachristos, E. (2001). Website evaluation: A usability-based perspective. In *Proceedings of the 8th Panhellenic Conference on Informatics*, 108–117.

Bonk, C. J. (2002). Online training in an online world. *Education at a Distance, 16*, 3. Retrieved from http://www.usdla.org/html/journal/MAR02_Issue/article02.html (last accessed Dec. 22, 2008).

Campbell, D. R., & Fiske, D. W. (1959). Convergent and discriminant validation by multi-trait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81–105.

Chang, Y. (2002). Assessing the usability of MOA interface designs. In *Proceedings of eLearn 2002*, 189–195.

Chin, J. P., Diehl, V. A., & Norman, K. L. (1988). Development of an instrument measuring user satisfaction of the human–computer interface. *Proceedings of CHI'88: Human Factors and Computing Systems*, 213–218.

Clark, R. C., & Mayer, R. E. (2003). *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning*. San Francisco: Jossey-Bass/Pfeiffer.

Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety: The experience of play in work and games*. San Francisco: Jossey-Bass.

Csikszentmihaly, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper & Row.

De Villiers, M. R. (2004). Usability evaluation of an e-learning tutorial: Criteria, questions and case study. G. Marsden, P. Kotze, A. Adesina-Ojo, Eds. In Proceedings of SAICSIT: Fulfilling the promise of ICT. ACM International Proceedings Series. pp. 284–291.

Feldstein, M. (2002). What is "usable" e-learning? *ACM eLearn Magazine,* 2002(9).

Gaver, W. W., & Martin, H. (2000). Alternatives. Exploring information appliances through conceptual design proposals. In *Proceedings of the CHI2000 Conference on Human Factors in Computing* (pp. 209–216). New York: ACM.

Hair, J. F., Anderson, R. E., Tatham, R. L. T., & Black, W. C. (1998). *Multivariate data analysis*. Upper Saddle River, NJ: Prentice-Hall.

Harun, M. H. (2002). Integrating e-learning into the workplace. *Internet and Higher Education*, *4*, 301–310.

Hassenzahl, M. (2003). The thing and I: Understanding the relationship between the user and product. In M. Blythe, C. Overbeeke, A. F. Monk, & P. C. Wright (Eds.), *Funology: From usability to enjoyment.* Dordrecht: Kluwer Academic Publishers. pp. 31–42.

Hassenzahl, M., & Tractinsky, N. (2006). User experience—A research agenda [Editorial]. *Behavior & Information Technology*, *25*(2), 91–97.

Hodges, C. B. (2004). Designing to motivate: Motivational techniques to incorporate in e-learning experiences. *The Journal of Interactive Online Learning 2*(3). Retrieved from http://www.ncolr.org/jiol/issues/PDF/2.3.1.pdf (last accessed Dec. 22, 2008)

Hornbaek, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human–Computer Studies*, *64*(2), 79–102.

Horton, W. (2000). *Designing Web-based training*. USA: John Wiley & Sons.
Hudlicka, E. (2003). To feel or not to feel: The role of affect in human–computer interaction. *International Journal of Human–Computer Studies*, *59*, 1–32.
Ingleton, C., & O'Regan, K. (1998). Recounting mathematical experiences: Exploring the development of confidence in mathematics. In *Proceedings of the AARE (Australasian Association for Research in Education) Conference.* Adelaide, SA: AARE Inc.
John, B. E., & Marks, S. J. (1997). Tracking the effectiveness of usability evaluation methods. *Behaviour & Information Technology*, *16* (4/5), 188–202.
Johnson, S., & Aragon, S. (2002). An instructional strategy framework for online learning environments. In *Proceedings of the World Conference on E-Learning in Corp., Govt., Health., & Higher Ed.*, 529–536.
Keeker, K. (1997). Improving web-site usability and appeal: Guidelines compiled by MSN usability research. http://msdn.microsoft.com/library/default.asp?url/=/library/en-us/dnsiteplan/html/improvingsiteusa.asp
Keinonen, T. (1998). *One-dimensional usability—Influence of usability on consumers' product preference*. Saarijärvi, Finland: Gummerus.
Keller, J. M. (1983). Motivational design of instruction. In C. M. Reigeluth (Ed.), *Instructional design theories and models: An overview of their current status*. Hillsdale, NJ: Erlbaum.
Keller, J. M., & Keller, B. H. (1989). *Motivational delivery checklist*. Florida State University.
Kirakowski, J., Claridge, N., & Whitehead, R. (1998). Human centered measures of success in Web site design. *Proceedings of Our Global Community*. Retrieved from http://www.research.att.com/conf/hfweb/proceedings/
Kirakowski, J., & Corbett, M. (1990). *Effective methodology for the study of HCI*. Amsterdam: Elsevier.
Konradt, U., & Sulz, K. (2001). The experience of flow in Interacting with a hypermedia learning environment. *Journal of Educational Multimedia and Hypermedia*, *10*(1), 69–84.
Konradt, U., Wandke, H., Balazs, B., & Christophersen, T. (2003). Usability in online shops: scale construction, validation and influence on the buyers intention and decision. *Behaviour & Information Technology*, *22*(3), 61–73.
Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human–Computer Interaction*, *7*(1), 57–78.
Lynch, P. J., & Horton, S. (1999). *Interface design for WWW Web style guide. Yale Style Manual*. Retrieved from http://info.med.yale.edu/caim/manual/interface.html (last accessed Dec. 22, 2008)
Mankoff, J., Dey, A. K., Hsieh, G., Kientz, J., Ames, M., & Ledered, S. (2003). Heuristic evaluation of ambient displays. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (pp. 169–176). New York: ACM Press.
Matsubara, Y., & Nagamachi, M. (1996). Motivation systems and motivation models for intelligent tutoring. In C. Frasson, G. Gauthier, A. Lesgold (Eds.), *Proceedings of the Third International Conference in Intelligent Tutoring Systems* (pp. 139–147). Berlin: Springer-Verlag.
Miller, R. H. (2002). E-learning Site Usability Checklist. Society for Technical Communication. Online archive available at: http://www.stcsig.org/usability/resources/toolkit/toolkit.html (last accessed Dec. 22, 2008).
Monk, A. (2002). Noddy's guide to usability. *Interfaces*, *50*, 31–33.
Moshinskie, J. (2001). How to keep e-learners from e-scaping. *Performance Improvement*, *40*(6), 28–35.
Nielsen, J. (2000). *Designing Web usability: The practice of simplicity*. New York: New Riders.
Nokelainen, P. (2006). An empirical assessment of pedagogical usability criteria for digital learning material with elementary school students. *Journal of Educational Technology & Society*, *9*(2), 178–197.

Notess, M. (2001). Usability, user experience, and learner experience. Retrieved from http://www.elearnmag.org/ (last accessed Dec. 22, 2008)

O'Regan, K. (2003). Emotion and e-learning. *Journal of Asynchronous Learning Networks*, *7*(3), 78–92.

Parlangeli, O., Marchigianni, E., & Bagnara, S. (1999). Multimedia system in distance education: Effects on usability. *Interacting with Computers*, *12*, 37–49.

Picard, R., Kort, B., & Reily, R. (2001). *Exploring the role of emotion in propelling the SMET learning process*. Retreived from http://affect.media.mit.edu/AC_research/lc/nsf1.PDF (last accessed Dec. 22, 2008)

Picard, R. W., & Daily, S. B. (2005). Evaluating affective interactions: Alternatives to asking what users feel. *CHI Workshop on Evaluating Affective Interfaces: Innovative Approaches.* Retrieved from http://www.sics.se/~kia/evaluating_affective_interfaces/Picard.pdf (last accessed Dec. 22, 2008).

Powell, T. A. (2000). Web Design: The Complete Reference. California, USA: McGraw Hill.

Quinn, C. N. (1996). Pragmatic evaluation: Lessons from usability. 13th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education. Adelaide: Australasian Society for Computers in Learning in Tertiary Education.

Reeves, T. C., Benson, L., Elliott, D., Grant, M., Holschuh, D., Kim, B., et al. (2002). Usability and instructional design heuristics for e learning evaluation. In *Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications*, 1615–1621.

Root, R. W., & Draper, S. (1983). Questionnaires as a software evaluation tool. In *Proceedings of the CHI 83* (pp. 83–87). New York: ACM.

Rust, J., & Golombok, S. (1989). *Modern psychometrics: The science of psychological assessment*. London: Routledge.

Ryu, Y. S., & Smith-Jackson, T. L. (2006). Reliability and validity of Mobile Phone Usability Questionnaire (MPUQ). *Journal of Usability Studies*, *2*(1), 39–53.

Saw, S. M., & Ng, T. P. (2001). The design and assessment of questionnaires in clinical research. *Singapore Medical Journal*, *42*(3), 131–135.

Schunk, D. (2000). *Learning theories: An educational perspective* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall.

Shneiderman, B. (1987). *Designing the user interface: Strategies for effective human–computer interaction*. Reading, MA: Addison-Wesley.

Smulders, D. (2002). Designing for learners, designing for users. *ACM eLearn Magazine*. Retrieved from http://www.elearnmag.org/(last accessed May, 2007)

Soloway, E., Guzdial, M., & Hay, K. (1994). Learner-centered design. The challenge for HCI in the 21st century. *interactions*, *1*(2), 36–48.

Spitzer. D. R. 1996, Motivation: The neglected factor in instructional design. Educational Technology, 36(3), 45–49.

Squires, D. (1999). Usability and educational software design: Special issue of *Interacting with Computers*. *Interacting with Computers*, *11*(5), 463–466.

Squires, D., & Preece, J. (1999). Predicting quality in educational software: Evaluating for learning, usability and the synergy between them. *Interacting with Computers*, *11*(5), 467–483.

Thomas, J. C., & Kellog, W. A. (1989). Minimizing ecological gaps in interface design. *IEEE Software*, *6*(1), 78–86.

Tselios, N., Avouris, N., Dimitracopoulou, A., & Daskalaki, S. (2001). Evaluation of distance-learning environments: Impact of usability on student performance. *International Journal of Educational Telecommunications*, *7*(4), 355–378.

Wade, V. P., & Lyng, M. (1999). Experience Improving WWW based courseware through usability studies. In *Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications*, 951–956.

Wang, Y. S. (2003). Assessment of learner satisfaction with asynchronous electronic learning systems. *Information & Management*, *41*, 75–86.

Weston, C., Gandell, T., McApline, L., & Filkenstein, A. (1999). Designing instruction for the context of online learning. *The Internet and Higher Education*, *2*(1), 35–44.
Whitelock, D., & Scanlon, E. (1996). Motivation, media and motion: Reviewing a computer supported collaborative learning experience. In P. Brna et al. (Eds.), *Proceedings of the European Conference on Artificial Intelligence in Education*.
Wolf, C. G., Caroll, J. M., Landauer, T. K., John, B. E., & Whiteside, J. (1989). The role of laboratory experiments in HCI: Help, hindrance, or ho-hum? *ACM SIGCHI Bulletin*, *20*, 265–268.
Zaharias, P. (2004). Usability and e-learning: The road towards integration. *ACM eLearn Magazine*, *6*.
Zaharias, P. (2005). E-learning design quality: A holistic conceptual framework. In Howard, Caroline, J. Boettcher, L. Justice, K. Schenk, P. L. Rogers, G. A. Berg, (Eds.), *Encyclopedia of distance learning*. New York: Idea Group Inc.
Zaharias, P. (2008). Usability in the context of e-learning: A framework augmenting "traditional" usability constructs with instructional design and motivation to learn. *International Journal of Technology and Human Interaction*. (in Review)
Zaharias, P., & Poulymenakou, A. (2006). Implementing learner-centered design: The interplay between usability and instructional design practices. *Journal of Interactive Technology and Smart Education*, *3*(2), 87–100.

APPENDIX

Rotated Component Matrix—8-Factor Solution

	Factor loadings (>.32)							
Items	*Factor 1*	*Factor 2*	*Factor 3*	*Factor 4*	*Factor 5*	*Factor 6*	*Factor 7*	*Factor 8*
I4	.690					.391		
CR1_5	.637				.388			
CR4_6	.589	.409						
C1	.587					.328		
CR8	.575			.323				
M1_2_3	.571						.333	
CR2	.559				.426			
I5	.517	.325					.429	
I1	.513							
CR7	.414			.328	.332			
CR3	.403	.341					.331	
F1		.774						
F2		.709						
LG4		.705					.339	
IA2		.673						
I3	.326	.514				.332		
IA1_3		.491		.388		.329		
N1_3_4			.712					
N6			.708					
N5			.703					
N2			.686					.359

(*Continued*)

Appendix (*Continued*)

	Factor loadings (>.32)							
Items	*Factor 1*	*Factor 2*	*Factor 3*	*Factor 4*	*Factor 5*	*Factor 6*	*Factor 7*	*Factor 8*
C2_3			.329					
A3				.659				
VD3				.616				
VD4	.337			.603				
LG1_2		.397		.599				
VD1_2	.327		.449	.461				
A2			.399	.416			.320	
CR9				.340	.652			
CR10					.649			
I2	.342	.349			.604			
L4					.502			
L1	.361				.408			.325
LD1						.777		
LD3						.696		
LD4			.346			.669		
LD2	.332				.344	.502		
A4							.731	
LG3				.329	.339		.527	
A1					.423		.516	
L2								.738
L3				.343				.616

Usability Items and Sources

Usability items used in questionnaire	*Literature*
Navigation	
Learners can decide what parts of the course to access, the order, and the pace	(Horton, 2000)
Learners can control their learning activities	(Reeves et al., 2002)
Learning units and modules are self-contained enough that learners can take them out of sequence without becoming confused	(Horton, 2000)
Navigation and access mechanisms (menus, course maps, indexes, etc.) are sufficient that learners can find specific items of content	(Horton, 2000)
The Web course allows the learner to leave whenever desired but easily return to the closest logical point in the course	(Reeves et al., 2000)
Learners always know where they are in the course	(Horton, 2000; Powell, 2000)
Learnability	
The layout of the wWeb course speaks for itself so that the extensive consultation of online help or other documentation does not interfere with learning and/or finding the desired information	(Reeves et al., 2000)

(*Continued*)

Appendix (*Continued*)

Usability items used in questionnaire	*Literature*
Learners can get started taking the Web course (locate it, install plug-ins, register, access starting page) using only online assistance	(Horton, 2000)
It is clear what learners should do if they get stuck or have questions	(Horton, 2000)
Learners can predict the general result of clicking on each button or link	(Horton, 2000)
Accessibility	
The pages and other components of the Web course download quickly	(Horton, 2000; Nielsen, 2000)
The Web course is easy to install, uninstall, and launch	(Chua, 2002)
The Web course includes information on technical requirements	(Chua, 2002)
The Web course is free from technical problems	(Chua, 2002)
Accommodation is made for users who turn off graphic information displays in their browser. Absence of graphics does not dilute the value of the information	(Miller, 2002)
Pages are displayed in such a way that those who choose to change default display characteristics to accommodate visual impairments or preferences may do so without loss of information	(Miller, 2002)
Consistency	
Terminology of the functions is used consistently throughout the Web course	(Miller, 2002)
The fonts, colors, and sizes are consistent throughout the Web course	(Chua, 2002)
The Web course maintains an appropriate level of consistency in its design from one part of the course to another	(Powell, 2000; Reeves et al., 2002)
Visual design	
The most important information on the screen is placed in areas most likely to attract the learner's attention	(Miller, 2002; Powell, 2000; Reeves et al., 2002)
The course is attractive and appealing to the learner's senses	(Chua, 2002; Keeker, 1997)
Text and graphics are legible	(Chua, 2002; Nielsen, 2000; Powell, 2000)
Fonts (style, color, saturation) are easy to read in both on-screen and printed versions	(Miller, 2002; Powell, 2000)
Interactivity	
The Web course does not provide too many long sections of text to read without meaningful interactions	(Reeves et al., 2002)
Learners are given the opportunity to practice ideas and skills immediately after they are presented	(Horton, 2000)
Practice activities exercise knowledge and skills in a way that prepares learners to apply learning to their jobs	(Horton, 2001)

(*Continued*)

Appendix (*Continued*)

Usability items used in questionnaire	*Literature*
The Web course engages learners in tasks that are closely aligned with the learning goals and objectives	(Chua, 2002)
The Web course uses games, simulations, role-playing activities, and case studies to gain the attention, sustain the interest, and maintain motivation of learner	(Chua, 2002; Keeker, 1997)
Content and recourses	
Content is organized in an appropriate sequence and in small modules for flexible learning	(Chua, 2002)
The material in the course is accurate and current	(Horton, 2000; Weston et al., 1999)
The course covers the subject in sufficient breadth and depth to meet the learning objectives	(Chua, 2002; Horton, 2001; Keeker, 1997)
Resources are provided in a manner that replicates as closely as possible their availability and use in the real world	(Reeves et al., 2002)
Text blocks are written in minimalist style: compact, yet useful	(Miller, 2002)
The Web course provides access to a range of resources (Web links, case studies, simulations, problems, examples) appropriate to the learning context	(Chua, 2002; Reeves et al., 2002)
Vocabulary and terminology used are appropriate for the learners	(Weston et al., 1999)
Abstract concepts (principles, formulas, rules, etc.) are illustrated with concrete, specific examples	(Horton, 2001)
All units/modules in the Web course include an overview and a summary	(Chua, 2002; Miller, 2002; Weston et al., 1999)
Learners are made aware of learning objective for each unit/module of the Web course.	(Chua, 2002; Miller, 2002)
Media use	
Graphics and multimedia assist in noticing and learning critical content rather than merely entertaining or possibly distracting learners	(Horton, 2001; Nielsen, 2000)
Graphics (illustrations, photographs, graphs, diagrams, etc.) are used appropriately; for example, to communicate visual and spatial concept	(Keeker, 1997; Nielsen, 2000)
Media (text, images, animations, etc.) included have a strong connection to the objectives and design of the courses.	(Horton, 2001; Keeker, 1997; Nielsen, 2000)
Learning strategies design	
The Web course provides opportunities and support for learning through interaction with others through discussion or other collaborative activities	(Quinn, 1996)
It is clear to the learner what is to be accomplished and what will be gained from its use	(Quinn, 1996)
The Web course is designed with activities that are both individual and group based	(Johnson & Aragon, 2002)
The Web course provides the learners opportunities to reflect	(Johnson & Aragon, 2002)

(*Continued*)

Appendix (*Continued*)

Usability items used in questionnaire	*Literature*
Instructional feedback	
The Web course provides learners with opportunities to access extended feedback from instructors, experts, peers, or others through E-mail or other Internet communications	(Reeves et al., 2002)
Feedback given at any specific time is tailored to the content being studied, problem being solved, or task being completed by the learner	(Reeves et al., 2002)
Instructional assessment	
The Web course provides opportunities for self-assessments that advance learner achievement	(Reeves et al., 2002)
Wherever appropriate, higher order assessments are (analysis, synthesis, and evaluation) provided rather than lower order assessments (recall and recognition)	(Reeves et al., 2002)
Posttests and other assessments adequately measure accomplishment of the learning objectives	(Horton, 2000; Quinn, 1996)
Learner guidance and support	
The online help or documentation is written clearly	(Reeves et al., 2002)
The online help is screen or context specific	(Albion, 1999; Reeves et al., 2002)
The Web course offers tools (taking notes, job aids, recourses, glossary, etc.) that support learning	(Chua, 2002)
The Web course provides support for learner activities to allow working within existing competence while encountering meaningful chunks of knowledge	(Quinn, 1996)