An Evaluation Instrument for Hypermedia Courseware

Georgiadou Elissavet
Researcher, University of Macedonia
Papakiriazi 12, Thessaloniki, 54645, Greece
elisag@otenet.gr

Anastasios A. Economides
Ass. Professor, University of Macedonia
Egnatias 156, Thessaloniki 54006, Greece
economid@uom.gr

Abstract

This paper presents an evaluation instrument for hypermedia courseware. Its design is based on an evaluation framework developed from the integration of a number of important issues emerged from research on instructional design and system evaluation the past fifteen years and is concerned with both social and practical acceptability of hypermedia courseware. The term social acceptability is related with the social basis of an educational system. Practical acceptability is examined through the evaluation of the following sectors: content, presentation and organization of the content, technical support and update processes and finally, the evaluation of learning. All sectors are equally important, as hypermedia courseware has to be simultaneously pedagogically and technically sound. The paper first discusses other evaluation efforts; next it introduces the evaluation framework and finally presents the evaluation instrument and suggests ways for the analysis of the results.

Keywords

Evaluation criteria, Evaluation instrument, Hypermedia courseware

Introduction

The number of products from the educational software industry has significantly increased the last decade and in particular there are numerous hypermedia courseware available in the market on almost any educational subject. (Courseware is a relatively recent appellation for Computer Based Learning, which refers to the use of computers for the delivery of instruction in an interactive mode.) The reason for that increase is closely related to the basic policy assumption that the educational system should serve the overall target of 'information society for all'. Schools should prepare students to use actively new information and communication technologies (ICTs) taking advantage of the life-learning process that these technologies support. As a result in most countries all curriculum are under continuous development adopting ICTs in teaching and learning.

Nowadays, hypermedia systems provide the necessary technology for highly interactive and potentially adaptive learning environments. Yet, in many cases authors of educational hypermedia are often tempted to impress rather than educate the user. As often stated, the failure of so many instructional programs has been the result of an emphasis solely on content, with little regard for principles of instructional design to produce effective, efficient, and appealing instruction. If hypermedia is not well designed, they will create difficulties for users, such as memory overload and divided attention, or they will fail to suit the variety of ways that people work together or alone (Preece, 1993).

The media and learning debate has carried on for several decades. In 1983, Richard Clark reviewed the research to that date on media-delivered education and concluded that instructional designers gain no learning benefits from employing a specific medium to deliver instruction (Clark, 1983). He claimed that any performance or time saving gains researchers observe is the result of uncontrolled instructional methods or novelty. In 1994 re-addressed the conclusions from his 1983 work by reviewing more recent studies (Clark, 1994). Kozma responded to Clark by arguing that media have an important role in learning as they can provide certain representations or model cognitive operations that are salient to a learning task, often ones that learners cannot or do not perform for themselves (Kozma, 1991). Some students will learn a task regardless of the delivery device. For others, though, Kozma argues that a careful use of media will enable learners to take advantage of its strengths to construct knowledge.

Elissavet, G., & Economides, A. A. (2003). An Evaluation Instrument for Hypermedia Courseware. *Educational Technology & Society*, 6(2), 31-44, Available at http://ifets.ieee.org/periodical/6-2/4.html

Despite the intricacies of the debate, in the early 1980s, several meta-analyses related to the effects of computers on learning were published by Kulik and his associates at the University of Michigan (Kulik, Kulik & Cohen, 1980; Kulik & Kulik, 1987; Kulik & Kulik, 1991) which has proven that computer-based instruction made small but significant contributions to the course achievement of all level students. Moreover, a meta-analysis on the effects of hypermedia versus traditional instruction on student's achievement on thirty five different empirical studies published from 1986 to 1997 showed that the effects of using hypermedia in instruction are positive and greater than the traditional instruction (Liao, 1997). Taking also into consideration cost-effectiveness and access issues regarding hypermedia versus traditional instruction, it can be argued that hypermedia courseware can be seen as an effective learning tool. However, instructional designers must carefully design hypermedia contents to take full advantage of them. Moreover, it can be argued that one of the main reasons for the lack of high quality of hypermedia courseware is that often research cannot keep pace with the advances of technology and as a result existing evaluation methods are often inadequate. Therefore, the development of evaluation criteria is very important for employing hypermedia courseware to best effect.

Background

Systematic evaluation of computer-based education (CBE) in all its various forms often falls behind development efforts (Flagg, 1990). There are several reasons for this lack of evaluation. Producers of CBE products are often invest more money in marketing CBE than in evaluating them. Moreover, consumers of technological innovations for education seem to assume that because these innovations are advertised as effective, they are effective. Also, evaluation of CBE has often been reduced to a number of indicators wherein the value of CBE is represented by the amount of money spent on hardware and software, the ratio of students to computers etc. (Becker, 1992). Another reason for the lack of the evaluation of CBE is the inadequate utility of the evaluations that have been previously conducted. Evaluation reports are usually presented in the format of social science research reports, "format that is almost useless for most educators" (Scriven, 1993).

However, besides the general trend, there are some important evaluation studies that either focus only on interface design or there are broader and focus on the pedagogical value of hypermedia systems as well. For example, heuristic evaluation suggested by Nielsen (Nielsen and Molich, 1990; Nielsen 1994) is looking at the usability problems in a user interface while Reeves (Reeves, 1992; Reeves and Harmon, 1994) pedagogical dimensions are used as criteria for evaluating different forms of computer-based education.

Heuristic evaluation is a usability engineering method for finding the usability problems in a user interface design so that they can be attended to as part of an iterative design process. Heuristic evaluation involves having a small set of evaluators examine the interface and judge its compliance with recognized usability principles (the 'heuristics'). The ten Usability Heuristics defined by Nielsen are: Visibility of system status; Match between system and the real world; User control and freedom; Consistency and standards; Error prevention; Recognition rather than recall; Flexibility and efficiency of use; Aesthetic and minimalist design; Help users recognize, diagnose, and recover from errors; Help and documentation.

Reeves, on the other hand, proposes fourteen pedagogical dimensions of computer-based education that can be used to compare one form of CBE with another or to compare different implementations of the same form of CBE. Each dimension is based on some aspect of learning theory or learning concept that can be used as criteria for evaluating different forms of computer-based education. These pedagogical dimensions are as follows: epistemology; pedagogical philosophy; underlying psychology; goal orientation; experiential value; teacher role; program flexibility; value of errors; motivation; accommodation of individual differences; learner control; user activity; cooperative learning; cultural sensitivity (Reeves, 1992; Reeves & Harmon, 1994).

Another important study is the 'conversational framework' for use in the analysis of teaching media developed by Laurillard (1993). Laurillard suggests that teaching media can be divided into four categories: discursive, adaptive, interactive and reflective. Discursive media should allow student and teacher to exchange views freely. Students must be able to act on, generate and receive feedback appropriate to the topic goal whilst the teacher must be able to reflect upon the student's actions and descriptions in order to adjust their own descriptions, making them more accessible for the student. Adaptive media allows the teacher to use the relation between his/her and the student's understanding to determine topic goals for the continuing session. Interactive media enables students, acting to achieve topic goals, to receive meaningful intrinsic feedback. Finally, reflective media facilitates teacher support for the process by which students link feedback on their actions to the topic goal. This list of required media characteristics was designed by Laurillard to encompass a complete specification of what is required of a learning situation. The 'conversational framework' incorporates all four categories of media.

Adaptation and reflection are internal to both teacher and student. The two levels in their dialogue: discursive and interactive are external processes transmitted over the media. Several other studies are based on the 'conversational framework' such as a framework for pedagogical evaluation of virtual learning environments (Britain & Liber, 1999) and evaluation of computer-supported collaborative learning (Crawley, 2002).

Besides the above evaluation studies there are also studies with a narrower centre of attention. For example the South Carolina Statewide Systemic Initiative developed an evaluation instrument for instructional material in mathematics; the Children's Software Review (1998) developed an evaluation instrument for children's Internet site; Biner (2002) created a web course evaluation questionnaire; Southern Regional Educational Board (2002) developed criteria for evaluating computer courseware; Beaudin and Quick (1996) formed an evaluation instrument for instructional video; etc.

It is also worth mentioning the standardization efforts of several organizations, such as the IMS Global Learning Consortium (IMS, 2001) and the International Standardization Organisation (ISO). IMS focus on interoperability - defining the technical specifications and supporting the incorporation of specifications into products and services worldwide. The ISO set up an ISO/IEC JTC1 SC36 sub-committee (ISO/IEC JTC1 SC36, 2001) "Standards for Information Technology for Learning, Education and Training". Again the focus here is interoperability and reusability of resources and tools. In addition, Europe has a number of bodies aiming on standardization of learning resources, such as the Alliance of Remote Instructional Authoring and Distribution Networks for Europe Foundation (ARIADNE, 2001) and the European Committee for Standardization: Information Society Standardization System (CEN/ISSS, 2001).

As shown, despite the fact that systematic evaluation of computer-based education (CBE) often falls behind development efforts there are several evaluation studies. However, some of the evaluation models described above require background knowledge on instructional technology, while the latter references have a very specific target (e.g. instructional material for mathematics, for children etc.). However, with the growth in the use of learning technologies and the availability of hypermedia courseware, an increasing number of teachers with no particular knowledge on instructional technology want to use such courseware in their teaching. The authors of this paper are attempting to provide an evaluation instrument for hypermedia courseware based on an evaluation framework, that can also address teachers with no particular knowledge on instructional technology, as a structured way of assisting them to initially assess a new piece of courseware that want to use in their teaching. Next, the evaluation framework is discussed.

Evaluation Framework

The efficiency of hypermedia courseware depends on many issues. In order to build the evaluation instrument the authors attempted to integrate in a framework a number of important issues emerged from research on instructional design and system evaluation the past fifteen years, and which should be considered from evaluators of hypermedia courseware (H.C.) that delivers mainly content knowledge (Georgiadou & Economides, 2000). It has to be acknowledged that this framework is relatively limited as there are numerous articles in the literature on instructional design and system evaluation. However, the authors, in order to develop the framework, tried to review a large number of them and then to focus on the most often cited authors and articles. Moreover, the framework is not rigid and therefore new parts could be added or existing ones could be altered as research advances in the area of educational hypermedia.

This framework is concerned with both social and practical acceptability of hypermedia courseware, based on Nielsen's idea that "the overall acceptability of a computer system is a combination of its social and practical acceptability" (Nielsen, 1990). The term social acceptability is related with the social basis of an educational system. In cases when the basis is teacher-centred, then the software that provides high levels of learner control and undermines the teacher's authority is possibly socially unacceptable. On the other hand, when the basis is student-centred, then a courseware that limits the student's potential for independent discovery is socially unacceptable. Moreover, as an example, whereas constructivist pedagogy advocates persistent questioning on the part of learners, questions, especially 'why?' questions, are inappropriate in cultures such as the Torres Strait Islanders of Australia. Although computer-based education may not be able to adapt to every cultural norm, they should be designed to be as culturally sensitive as possible (Powell, 1993).

Given that a piece of hypermedia courseware is socially acceptable, its practical acceptability is examined through the evaluation of the following four sectors: a) content, b) presentation and organization of the content, c) technical support and update processes and finally, d) the evaluation of learning. All sectors are equally important, as hypermedia courseware has to be simultaneously pedagogically and technically sound. Moreover, each sector includes a number of criteria that are incorporated in the evaluation instrument, which should be met in a satisfactory level, in order to characterize a piece of hypermedia courseware of high quality. Furthermore, cost-effectiveness should always be examined when similar products seem to have the same educational value. Figure 1 presents in a diagram the sectors included in the framework and the factors that are associated with them.

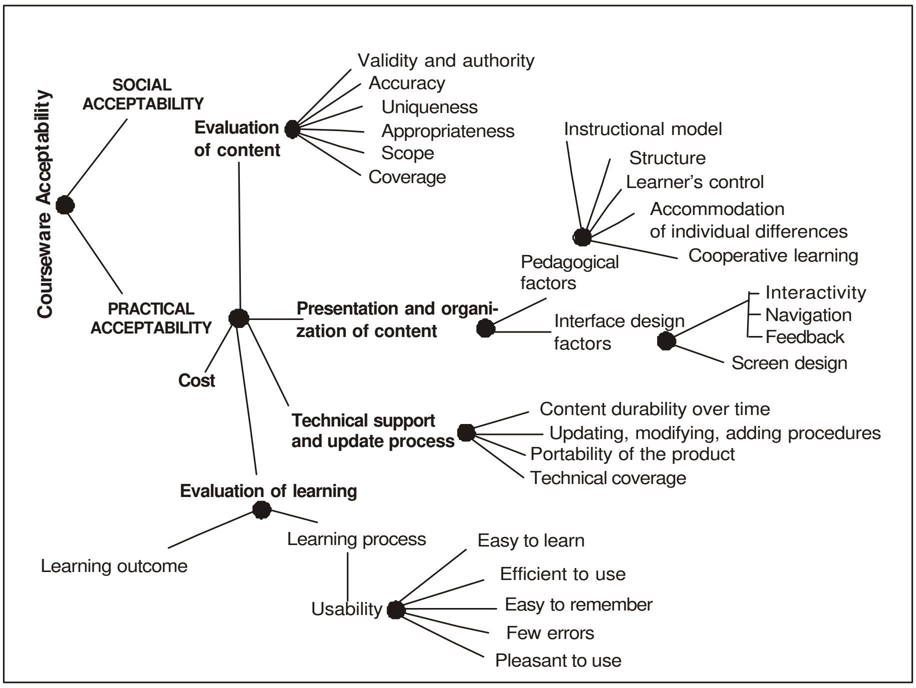

Figure 1. Diagram of the Evaluation Framework

Before we proceed to present the evaluation instrument it is necessary to discuss the underlying theory of the criteria used for the 'presentation and organization of the content' and the 'evaluation of learning' sectors. It is not necessary to do the same regarding the criteria for the evaluation of 'content' and those for the 'technical support and update processes' as the relevant items included in the evaluation instrument that are based on guidelines available in the literature (ANSI Standards Committee on Dental Informatics, 2002; Southern Regional Educational Board, 2002) are self-explanatory.

Presentation and Organization of the Content

The factors associated with this sector are the pedagogical ones that are concerned with learning and instructional design theories and the interface design factor.

Pedagogical Factor: This is a complicated factor as there are different beliefs of how humans learn. However, cognitive theories stress that learning is an active, constructive, cumulative, self-regulated process in which the learner plays a critical role. Moreover, current instructional theory focusing on learner-centred approaches depends on information access and learning environments that encourage free interaction with information. The agreement with the principles of an instructional design theory depends heavily on the subject matter. In addition, teachers' beliefs is of great importance, especially in cases when the hypermedia courseware is part of the curriculum.

Nevertheless, the two core elements that are important in all educational settings are 'motivation' and 'structure', which largely define the instructional nature of an information environment. A typical way to motivate the learner is to inform him/her what will s/he achieve at the end of the instruction by stating the aims and objectives (Gagné, Briggs, Wager, 1988). As far as the structure of the hypermedia courseware is concerned, that is how to organise instructional information, this again depends on the subject matter. However, in cases when the instructor wants to permit the learners to advance, review, see examples, repeat the unit, or escape to explore another unit, Jonassen (1992) suggest the network type or structured hypermedia as most appropriate. Structured hypermedia consists of sets of nodes, each set accessible from any other set. The node sets can be structured in any number of ways, such as node-link, hierarchical, network, depending on the nature of the processing the designer wants to elicit from the user. The structure of each node set with the various options available within each set needs to be conveyed on every screen. Another method for structuring the node sets is to combine related concepts, tie them together in an introductory block, and then permit access within the set only to concepts contained within the set.

In hypermedia learning systems another important element is ‘learner's control’, which is primary in the design of interactive learning as it allows students to tailor the learning experience to their own individual needs. However, there are dangers in surrendering too much control to the user, as low-ability students may get confused when control depends on a wide range of options (Gray, 1989; Litchfield, 1993). The high level of learner control may result in disorientation and distraction. The amount and type of learner control depends on the learner characteristics (age and cognitive capabilities), content, and the nature of the learning task (Poncelet & Proctor, 1993). Content that must be mastered and unfamiliar tasks often requires more program control, compared to content with no qualified mastery levels or familiar learning tasks. Learner control is more appropriate than program control when learners are more capable and are familiar with the learning task. Moreover, advisement is provided to assist learners in making decisions and control is used consistently within a lesson (Ross & Morrison, 1989). In general, the more control is given to the learners, the more feedback about their decisions should be given (Mcateer & Shaw, 1995).

Moreover, the issues of ‘accommodation of individual differences’, and ‘cooperative learning’ are highly important in the effectiveness of hypermedia-based learning. In most education contexts learners are not homogeneous in terms of prerequisite knowledge, motivation, experience, learning styles and cognitive styles. Also evidence suggests that when hypermedia-learning systems are structured to allow cooperation, learners benefit both instructionally and socially.

Interface Design Factor: *Interactivity - Navigation - Feedback*: Interactivity in instruction comprises the nature of the activity performed by the technology and the learner, as well as the ability of the technology to adapt the events of instruction in order to make that interaction more meaningful (Reigeluth, 1987). It is important to design as much meaningful interactivity as possible into instructional software (Orr, Golas, & Yao, 1994). The amount of navigational assistance needed is a function of the size of the knowledge base, the usefulness of navigational aids that are already part of the authoring software, and the types of links the software allows (Locatis, Letourneau & Banvard, 1989). Guidelines for increased interactivity were produced from researchers (Shneiderman & Kearsley, 1989; Tessmer, Jonassen & Caverly, 1989) and are used in the instrument as evaluation items in the relevant section.

The basic factors that can determine the effectiveness of feedback are the type and frequency of feedback given and the delay between feedback and instruction (Jonassen & Hannum, 1987). Feedback is closely related with the issue of interaction, as action without feedback is completely unproductive for a learner. Laurillard (1993) identifies two types of feedback, ‘intrinsic’ and ‘extrinsic’. *Intrinsic* feedback is what is given as a natural consequence of an action. To illustrate the concept of *intrinsic* feedback Laurillard uses examples of a child's actions while playing with water as the physical world responds to the child's actions of filling, pouring, etc. On the other hand *extrinsic* feedback does not occur within a situation but as an external comment on it: right or wrong. She suggests that extrinsic feedback is not a necessary consequence of the action, and therefore is not expressed in the world of the action itself, but at the level of the description of the action. In computer-based instruction, however, the *intrinsic* feedback relates to navigation and interactivity with the instructional program, and the *extrinsic* feedback relates to the feedback on user's performance. Schimmel (1988) identifies three types of extrinsic feedback: (a) Confirmation feedback that simply confirms whether a learner's answer is correct or incorrect; (b) Correct response feedback that presents the correct answer; (c) Explanatory feedback, such as a step-by-step solution to an incorrectly answered question. Many actions require more extended *extrinsic* feedback than confirmation feedback. Simple answers such as right or wrong cannot provide any information about how learners should correct their performance. A more helpful form of extrinsic feedback would give the learner information about how to adapt and correct their performance, such as correct response and explanation feedback.

‘**Screen design’** is also an important evaluation factor. Different screen elements should be used to present stimulating information that will motivate and assist the learners in retaining and recalling the information. The psychological limitations to consider when designing hypermedia learning systems include: (a) Memory load: i.e. how many different control icons is it reasonable for learners to remember at any one time? (b) Perception:

i.e. what colours and fonts provide the best readability?, and (c) Attention: i.e. how can the users' attention be drawn to information that is relevant, when there is a lot of different information on the screen? (Preece, 1993). A large number of screen design guidelines produced from several researchers on educational technology exist in the literature and the relevant items on the evaluation instrument are based on these (Morris, Owen & Fraser, 1994; Cox & Walker, 1993; Clarke, 1992; Mcateer & Shaw, 1995).

Evaluation Of Learning

Marchionini (1990) argued that the interactivity of hypermedia systems provides learners with access to vast amount of information in varied forms, control over the process of learning, and the potential for collaboration with the system and other people. Such empowerment of learners forces evaluators of learning to adopt a broad-based set of methods and criteria to accommodate 'self-directed' learning. He proposes a 'multi-faceted' approach to the evaluation of hypermedia based learning that address both the outcomes and the processes of learning.

The *learning outcomes* are evaluated through performance tests typically used to judge the quality and the quantity of learning, which usually have the form of 'pre-tests' used to determine learning outcomes prior to the intervention and 'immediate' and 'delayed post-tests' to examine learning outcomes after the intervention. The *learning process* refers to the usability of a product and should be evaluated by observing and measuring the end-users attitudes. Usability is usually associated with five parameters (Nielsen, 1990): (1) Easy to learn: Users can quickly get some work done with the system, (2) Efficient to use: Once the user has learnt the system, a high level of productivity is possible, (3) Easy to remember: The casual user is able to return to using the system after some period without having to learn everything all over, (4) Few errors: Users do not make many errors during the use of the system or if they do so they can easily recover them, and (5) Pleasant to use: Users are subjectively satisfied by using the system.

Evaluation Instrument

The criteria selected from the literature for every sector of the evaluation framework were used as the basis for the design of the initial version of the evaluation instrument. This initial version was disseminated for comments to academics, postgraduate students and researchers in the field of educational technology at the University of Macedonia, Greece. This effort was under a project run for two years (2000-2001) by the University of Macedonia called EPENDISI that aimed to train secondary schoolteachers in the use of ICTs in the classroom and also to build a database that contains information and resources on several evaluated educational software on almost all secondary school subjects. Taking into consideration the comments provided the instrument was revised and its final form is presented here.

The instrument has the form of a suitability scale questionnaire with five points; where figure (1) is assigned to strongly agree and figure (5) to strongly disagree. The scale also includes the figure (0) for those items in the questionnaire that cannot be evaluated, as they do not apply during the evaluation of particular hypermedia courseware. One hundred and twenty four items are included in the instrument and they cover both cases of stand-alone and web-based hypermedia courseware. The one hundred items refer to both stand-alone and web-based ones and the extra 24 items refer only to web-based ones, as these applications have some distinct characteristics regarding screen design and technical support and update processes. However, the instrument does not include items regarding the Social Acceptability because the criteria for such an evaluation cannot have universal application, as different educational systems have different beliefs on what is socially acceptable or unacceptable; therefore these criteria should be determined every time from the evaluators of each educational system.

The different sections of the instrument and the items included are presented next. It has to be noted at this point that the numeration of the items continues from each previous section in order to be more helpful to potential evaluators.

?. Evaluation of the content

1. The content is reliable	0 1 2 3 4 5
2. The origin of information is known	0 1 2 3 4 5

3. The authors and the publishers are reputable	0 1 2 3 4 5
4. Balanced presentation of information	0 1 2 3 4 5
5. Bias-free viewpoints and images	0 1 2 3 4 5
6. Balanced representation of cultural, ethnic and racial groups	0 1 2 3 4 5
7. Correct use of grammar	0 1 2 3 4 5
8. Current and error-free information	0 1 2 3 4 5
9. Concepts and vocabulary relevant to learners' abilities	0 1 2 3 4 5
10. Information relevant to age group curriculum	0 1 2 3 4 5
11. Information of sufficient scope and depth	0 1 2 3 4 5
12. Logical progression of topics	0 1 2 3 4 5
13. Variety of activities, with options for increasing complexity.	0 1 2 3 4 5

?. Organization and Presentation of the Content

?.1 Pedagogical Parameters

?.1.1. Instructional Theories – Curriculum

14. The design of the hypermedia courseware is based on reliable learning and instructional theories and is directly related with the content of the curriculum.	0 1 2 3 4 5
15. The application of the hypermedia courseware is possible in various topics of the curriculum	0 1 2 3 4 5
16. The application of the hypermedia courseware is possible on issues related with the curriculum	0 1 2 3 4 5
17. The hypermedia courseware can be used by learners alone, without the need of other instructional objects (i.e. book)	0 1 2 3 4 5

?.1.2. Structure

18. The content is structured in a clear and understandable manner	0 1 2 3 4 5
19. The structure allows learners to move around freely in different units	0 1 2 3 4 5
20. The structure of the H.C. permits learners to advance, review, see examples, repeat the unit, or escape to explore another unit	0 1 2 3 4 5

?.1.3. Learners Control

21. Learner's control corresponds to learners' age	0 1 2 3 4 5
22. Learner's control corresponds to learners' cognitive capabilities	0 1 2 3 4 5
23. The quantity of learner's control corresponds with the feedback given from the H.C.	0 1 2 3 4 5

?.1.4. Adaptivity

24. The H.C. considers the individual differences of the learners	0 1 2 3 4 5
25. The H.C. considers the different learning styles	0 1 2 3 4 5
26. The H.C. considers different background knowledge of the learners	0 1 2 3 4 5
27. The H.C. considers the different motivation of the learners	0 1 2 3 4 5
28. The H.C. considers the different learning experience	0 1 2 3 4 5
29. The H.C. contains understanding assignments	0 1 2 3 4 5
30. The H.C. contains assignments that develop the critical ability	0 1 2 3 4 5
31. The H.C. facilitates learning by doing	0 1 2 3 4 5
32. The H.C. permits learners to change the basic configurations	0 1 2 3 4 5
33. The H.C. offers the ability to change the level of difficulty	0 1 2 3 4 5
34. The H.C. allows learners to work on their own pace	0 1 2 3 4 5

?.1.5. Collaborative learning

35. The H.C. promotes collaborative learning	0 1 2 3 4 5
36. The H.C. contains assignments that can be executed by a group of learners	0 1 2 3 4 5
37. The H.C. encourages discussion and collaboration among learners	0 1 2 3 4 5

?2. Design Factors

?.2.1 Interactivity - Navigation - Feedback

?.2.1.1. Interactivity

38. The interactivity of the H.C. is according to the maturity of the students	0 1 2 3 4 5
39. The H.C. provides opportunities for interaction at least every three or four screens	0 1 2 3 4 5
40. The content is chunked into small segments and includes build in questions, reviews, and summaries for each segment	0 1 2 3 4 5
41. The H.C. poses frequently questions to the users that do not interrupt the learning process	0 1 2 3 4 5
42. The H.C. ask students to apply what they have learnt rather than memorise it	0 1 2 3 4 5
43. The H.C. uses rhetorical questions during instruction to get students to think the content	0 1 2 3 4 5
44. The H.C. allows learners to discover information through active exploration	0 1 2 3 4 5

?.2.1.2. Navigation

The H.C. includes:

45. Help key to get procedural information	0 1 2 3 4 5
46. Answer key for answering a question	0 1 2 3 4 5
47. Glossary key for seeing the definition of any term	0 1 2 3 4 5
48. Objective key for reviewing the course's objectives	0 1 2 3 4 5
49. Content map key for seeing a list of options available	0 1 2 3 4 5
50. Summary and review key for reviewing whole or parts of the lesson	0 1 2 3 4 5
51. Menu key for returning to the main page	0 1 2 3 4 5
52. Exit key, for exiting the program	0 1 2 3 4 5
53. Comment key for recording a learner's comment	0 1 2 3 4 5
54. Example key for seeing examples of an idea	0 1 2 3 4 5
55. Key for moving forward or backward in a lesson	0 1 2 3 4 5
56. Key for accessing the next lesson in a sequence	0 1 2 3 4 5

?.2.1.3. Feedback

57. The H.C. provides feedback immediately after a response	0 1 2 3 4 5
58. The placement of feedback is varied according to the level of objectives. (Provide feedback after each response for lower level objectives, and at the end of the session for the higher level ones)	0 1 2 3 4 5
59. The H.C. provides feedback to verify the correctness of a response	0 1 2 3 4 5
60. For incorrect responses, information is given to the student about how to correct their answers, or hints to try again	0 1 2 3 4 5
61. The H.C. allows students to print out their feedback	0 1 2 3 4 5
62. The H.C. allows students to check their performance	0 1 2 3 4 5
63. The H.C. allows students to measure the time they consume in a certain on-line assignment	0 1 2 3 4 5

?.2.2 Screen Design

64. Screens are designed in a clear and understandable manner	0 1 2 3 4 5
65. The presentation of information can captivate the attention of students	0 1 2 3 4 5
66. The presentation of information can stimulate recall	0 1 2 3 4 5
67. The design does not overload student's memory	0 1 2 3 4 5
68. The use of space is according to the principles of screen design	0 1 2 3 4 5
69. The design uses proper fonts in terms of style and size	0 1 2 3 4 5
70. The use of text follows the principles of readability	0 1 2 3 4 5
71. The color of the text follows the principles of readability	0 1 2 3 4 5
72. The number of colors in each screen is no more than six	0 1 2 3 4 5
73. There is consistency in the functional use of colors	0 1 2 3 4 5
74. The quality of the text, images, graphics and video is good	0 1 2 3 4 5
75. Presented pictures are relevant to the information included in the text	0 1 2 3 4 5
76. The use of graphics support meaningfully the text provided	0 1 2 3 4 5
77. A high contrast between graphics and background is retained.	0 1 2 3 4 5
78. There is only one moving image (animation and/or video) each time on the same screen	0 1 2 3 4 5
79. Video enhance the presentation of information	0 1 2 3 4 5
80. Sound is of good quality and enhances the presentation of information	0 1 2 3 4 5
81. Sound is an alternative means of presenting information and not a necessity (except for music and language courses)	0 1 2 3 4 5
82. The integration of presentation means is well coordinated	0 1 2 3 4 5

C. Technical Support and Update Process

83. The content has durability over time	0 1 2 3 4 5
84. The content can be updated and/or modified with new knowledge that will appear soon after the purchase of the courseware	0 1 2 3 4 5
85. Technical coverage is offered from the production company	0 1 2 3 4 5
86. The courseware can be used in different platforms	0 1 2 3 4 5
87. Documentation exist regarding technical requirements for software and hardware needed	0 1 2 3 4 5
88. There are instructions for the installation and use of the courseware	0 1 2 3 4 5
89. There is a review of the courseware's contents for use by the instructor	0 1 2 3 4 5
90. Documentation exists regarding the use of the courseware in the classroom with teaching plans and related activities	0 1 2 3 4 5
91. The updating, modifying and adding procedures are relatively easy for the average user	0 1 2 3 4 5
92. The H.C. provides printing capabilities	0 1 2 3 4 5
93. The H.C. allows to keep (save) every step of the activities	0 1 2 3 4 5

D. Evaluation of learning

D.1 The process of learning

94. The H.C. is easy to learn; the user can quickly get some work done with it	0 1 2 3 4 5
95. The H.C. is efficient to use; once the user has learnt it, a high level of productivity is possible	0 1 2 3 4 5
96. The H.C. is easy to remember; the casual user is able to return to using it after some period without having to learn everything all over	0 1 2 3 4 5
97. The structure of the H.C. is comprehensive and the average performance learners can easily follow it	0 1 2 3 4 5
98. Users do not make many errors during the use of the H.C. or if they do so they can easily recover them	0 1 2 3 4 5
99. Users are subjectively satisfied by using the H.C	0 1 2 3 4 5

100. Users find the H.C. interesting	0 1 2 3 4 5

In cases when the hypermedia courseware is web-based then additionally the following items are examined as well for the Screen Design section.

1. The speed of the program (download) is satisfactory	0 1 2 3 4 5
2. Horizontal scrolling bars are not used	0 1 2 3 4 5
3. The hypermedia courseware includes local links in order to facilitate navigation	0 1 2 3 4 5
4. ?he H.C. is flexible and allows students to access all its contents	0 1 2 3 4 5
5. The first page is understandable	0 1 2 3 4 5
6. The H.C. in general has a distinct and easily recognized character	0 1 2 3 4 5
7. The information is organized into small and functional units	0 1 2 3 4 5
8. The H.C. includes alternative ways of presentation (e.g. with or without graphics)	0 1 2 3 4 5
9. The H.C. includes content map	0 1 2 3 4 5
10. The H.C. includes search engine	0 1 2 3 4 5
11. The main navigation tools are always on display to increase speed of use and save from backtracking	0 1 2 3 4 5
12. The way that the navigation tools work is easily understandable from the students	0 1 2 3 4 5
13. Each learning unit is presented under the same design principles (consistency)	0 1 2 3 4 5
14. External links are loaded in a separate window	0 1 2 3 4 5
15. The H.C. includes synchronous communication channels	0 1 2 3 4 5
16. The H.C. includes asynchronous communication channels	0 1 2 3 4 5

Moreover, for web-based hypermedia courseware the following items need examination for Technical Support and Update Process section.

17. The H.C. includes information regarding how often is updated	0 1 2 3 4 5
18. The H.C. includes information regarding its latest update	0 1 2 3 4 5
19. The links are stable	0 1 2 3 4 5
20. The frequency of malfunction is rare	0 1 2 3 4 5
21. The courseware includes mirror sites	0 1 2 3 4 5
22. The content is updated regularly	0 1 2 3 4 5
23. The management and the maintenance of the site is satisfactory	0 1 2 3 4 5
24. The H.C. includes archives from previous editions	0 1 2 3 4 5

Evaluation Process and Analysis of the Results

As shown from the items included in the instrument, during the evaluation of a hypermedia courseware application a number of people should be involved, i.e. content experts, instructional technologists, educators and interface designers. However, the items are quite straightforward and as a result the instrument can be used from educators with no particular knowledge on instructional technology, as a structured way of assisting them during the initial evaluation of a new piece of courseware that want to use in their teaching. After this initial stage, an evaluation with the students is required in order for educators to have a better understanding of the courseware's value and potential.

In order to analyse the results the evaluators have to consider that not all the factors have the same weight; and content is the most important of all. If the content does not meet the educator's criteria then there is no need to further evaluate the organization and the presentation of the educational material. However, to have an overall idea regarding the value of the courseware at the end of the evaluation process for a particular courseware the sum of the score in all items - except those resulted from the evaluation of the content - and its comparison with the total sum, that is the maximum of the marks in all items is required. Therefore, by excluding the 13 items for the evaluation of content the total sum for stand-alone applications is 435 (87*5) and 555 (111*5) for web-based ones (Table 1). These two figures need alteration in the case that not all the items were used during the evaluation, as some of them could not find application in certain pieces of hypermedia courseware. For example, if only 80 items are used then the total sum is 400 (80*5).

Stand-alone		Web-based	
Total sum	*Score*	*Total sum*	*Score*
435		555	

Table 1. Assessment table for all the items of the evaluation instrument

When evaluating two or more courseware on the same subject, then the above figures can be a useful starting point in determining the most appropriate one. Yet, the most important part of the evaluation is the examination of the scores resulted from the evaluation of the four different sectors separately: a) content, b) presentation and organization of the content, c) technical support and update processes and finally, d) the evaluation of learning. The examination of these scores is important in order to secure the case that an application is technically sound but does not have a pedagogical value and vice versa. Table 2 can be used to compare the results.

	Assessment of the Different Sectors			
	Stand-alone		Web-based	
	Total sum	*Score*	*Total sum*	*Score*
A. Content	65		65	
?. Organisation and Presentation of the content	340		420	
B1. Pedagogical Parameters	120		120	
B2. Design factors	225		305	
C. Technical Support and Update Process	55		95	
D. Evaluation of learning	35		35	

Table 2. Assessment table for the different s ectors of the evaluation instrument

It has to be mentioned that in order to ensure high quality of hypermedia courseware the evaluators' team (or the teacher) potentially could agree on some standards and set a threshold to the comparison of the results. For example, if the score resulted from the evaluation of an application is not equal with the two thirds of the total sum in all sectors then the application cannot be used for teaching and learning.

Summary

This paper presented an evaluation instrument for hypermedia courseware that is designed according to an evaluation framework developed from the integration of a number of important issues emerged from research on instructional design and system evaluation the past fifteen years and is concerned with both social and practical acceptability of hypermedia courseware. One hundred and twenty four items are included in the instrument that has the form of a suitability scale questionnaire that are concerned with the evaluation of four main sectors: a) content, b) presentation and organization of the content, c) technical support and update processes and finally, d) the evaluation of learning.

Postgraduate students and secondary schoolteachers in the University of Macedonia, Greece, used the instrument during 2001, in order to evaluate hypermedia courseware on almost all secondary school subjects of the Greek curriculum. This effort was under a project run for two years (2000-2001) by the University of Macedonia, Greece called EPENDISI that aimed to train secondary schoolteachers in the use of ICTs in the classroom and also to build a database that contains information and resources on several evaluated educational software on secondary school subjects. During the evaluation period users of the instrument express their opinion on the instrument itself during debriefing sessions. In general, they agreed that it was easy to use as most of the items included are clear-cut and also the analysis of the results was a simple process that gives relatively quickly an overall idea of a particular courseware's value. Moreover, secondary schoolteachers stated that the first time they used the instrument they felt a bit frustrated as they had little knowledge on instructional design and they usually were consumers of the product rather than evaluators. However, after using the instrument for more than three times they had a better understanding of instructional design and system's evaluation and as a result they felt comfortable with the evaluation process. However, most of the instrument users stated that in order to determine the real value of a particular courseware evaluation with the end-users (i.e. students) is essential.

As research progresses in the field of hypermedia courseware evaluation new items can be added to the presented instrument. Therefore, it is a flexible tool that could be easily adapted in an educational environment and its improvement could be an ongoing process.

References

ANSI Standards Committee on Dental Informatics - Working Group Educational Software Systems (2002). *Guidelines for the design of educational software*, http://www.temple.edu/dentistry/di/edswstd/.

ARIADNE: Alliance of Remote Instructional Authoring and Distribution Networks for Europe, http://www.ariadne-eu.org/.

Beaudin, B. P., & Quick, D. (2002). Instructional Video Evaluation Instrument. *Extension Journal*, http://www.joe.org/joe/1996june/a1.html.

Becker, H. J. (1992). Computer education. In M. C. Alkin (Ed.) *Encyclopedia of educational research*, New York: Macmillan.

Biner, P. (2002). *Web Course Evaluation Questionnaire*, http://www.distance-educator.com/dnews/.

Britain, S., & Liber, O. (1999) *A Framework for Pedagogical Evaluation of Virtual Learning Environments.* Report 41, October, University of Wales: JISC Technology Application Programme.

CEN/ISSS *The European Committee for Standardization: Information Society Standardization System*, http://www.cenorm.be/isss/Workshop/lt/.

Children's Internet Site Evaluation Instrument (1988). *Children's Software Review*, January/February, http://www.childrenssoftware.com/contact.html.

Clark, R. E. (1994). Media will Never Influence Learning. *Educational Technology Research and Development*, 42 (2), 21-29.

Clark, R. E. (1983). Reconsidering Research on Learning from Media. *Review of Educational Research*, 53 (4), 445-459.

Clarke, A. (1992). *The Principles of Screen Design for Computer Based Learning Materials*. U.K.: Department of Employment.

Cox, K., & Walker, D. (1993). *User Interface Design* (2nd Ed.), New York London: Prentice Hall.

Crawley, R. M. (2002). *Evaluating CSCL - Theorists' & Users' Perspectives*, University of Brighton: Collaborative Computing Research Group, http://www.bton.ac.uk/cscl/jtap/paper1.htm.

Flagg, B. (1990). *Formative Evaluation for Educational Technologies*, Hilsdale, New Jersey: Lawrence Erlbaum Associates.

Gagné, R. M., Briggs, L. J., & Wager W. W. (1988). *Principles of Instructional Design*, N.Y.: Holt, Rinehart, and Winston.

Georgiadou, E., & Economides, A. (2000). Evaluation Factors of Educational Software. In Kinshuk, Chris Jesshope & Toshio Okamoto (Eds.) *Proceedings of the International Workshop on Advanced Learning Technologies Proceedings*, Los Alamitos, CA: IEEE Computer Society, 113-116.

Gray, S. H. (1989). The effect of locus of control and sequence control on computerised information retrieval and retention. *Journal of Educational Computing Research*, 5 (4), 459-471.

IMS Global Learning Consortium, http://www.imsproject.org/.

ISO/IEC JTC1 SC36, http://jtc1sc36.org/.

Jonassen, D. H. (1992). Designing Hypertext for Learning. In Scanlon, E., & O'Shea, T. (Eds.) *New Directions in Educational Technology*, Berlin: Springer-Verlag, 123-130.

Jonassen, D. H., Hannum, W. H. (1987). Research-based Principles for Designing Computer Software. *Educational Technology*, 1 (18), 42-51.

Kozma, R. (1991). Learning with Media. *Review of Educational Research,* 61 (2), 179-211.

Kulik, C-L., & Kulik, J. A. (1991). Effectiveness of computer-based instruction: an updated analysis. *Computers and Human Behavior*, 7, 75-94.

Kulik, J. A., & Kulik, C-L. (1987). Review of recent research literature on computer-based instruction. *Contemporary Educational Psychology*, 12, 222-30.

Kulik, J. A., Kulik, C., & Cohen, P. A. (1980). Effectiveness of Computer-based College Teaching: A Meta-Analysis of Findings. *Review of Educational Research.* 50 (4), 525-544.

Laurillard, D. (1993). *Rethinking University Teaching*, London: Routledge.

Liao, Y. C. (1997). Effects of Hypermedia vs. Traditional Instruction on Student Performance: A Meta-Analysis. *CD-ROM Proceedings of WebNet97 World Conference on the WWW, Internet, and Intranet.* Norfolk VA, USA: AACE, 312-317.

Litchfield, B. (1993). Design Factors in Multimedia Environments: Research Findings and Implications for Instructional Design. *Annual Meeting of the American Educational Research Association*, 1-10.

Locatis, C., Letourneau, G., & Banvard, R. (1989). Hypermedia and Instruction, *Educational Technology Research and Development*, 37 (4), 65-77.

Marchionini, G. (1990). Evaluating Hypermedia Based Learning. In Jonassen D., & Mandl H. (Eds.) *Designing Hypermedia for Learning*, Berlin: Springer-Verlag, 355-374.

Mcateer, E., & Shaw, R. (1995). *The Design of Multimedia Learning Programs*, University of Glasgow: EMASHE Group.

Morris, J. M., Owen G. S., & Fraser, M. D. (1994). Practical Issues in Multimedia User Interface Design for Computer-Based Instruction. In Resman, S. (Ed.) *Multimedia Computing: Preparing for the 21st Century*, London: Idea Group Publishing, 225-284.

Nielsen, J. (1990). Evaluating Hypertext Usability. In Jonassen, D., & Mandl, H. (Eds.) *Designing Hypermedia for Learning*, Berlin, Heidelberg: Springer-Verlag, 147-168.

Nielsen, J., & Molich, R. (1990). Improving a human-computer dialogue. *Communications of the ACM*, 33 (3), 338-348.

Nielsen, J. (1994). Heuristic evaluation. In Nielsen, J., & Mack, R. L. (Eds.) *Usability Inspection Methods*, New York: John Wiley & Sons, 25-64.

Orr, K. L, Golas, K. C., & Yao, K. (1994.) Storyboard Development for Interactive Multimedia Training. *Journal of Interactive Instruction Development*, Winter, 18-31.

Poncelet, G. M., & Proctor, L. F. (1993). Design and Development Factors in the Production of Hypermedia-based Courseware, *Canadian Journal of Educational Computing*, 22 (2), 91-111.

Powell, G. C. (1993). Incorporating learner cultural diversity into instructional systems design: An investigation of faculty awareness and teaching practices. *Unpublished doctoral dissertation*, The University of Georgia.

Preece, J. (1993). Hypermedia, Multimedia and Human Factors. In Latchem, C., Williamson, J., & Henderson-Lancett, L. (Eds.) *Interactive Multimedia*, London: Kogan Page, 135-149.

Reeves, T. C. (1992). Effective dimensions of interactive learning systems. *Invited keynote paper presented at the Information Technology for Training and Education Conference* (ITTE `92), Queensland, Australia.

Reeves, T. C., & Harmon, S. W. (1994). Systematic Evaluation Procedures for Interactive Multimedia for Education and Training. In Resman, S. (Ed.) *Multimedia Computing: Preparing for the 21st Century*, Harrisburg, London: Idea Group Publishing, 472-505.

Reigeluth, C. M. (1987). *Instructional Theories in Action*, Hilsdale, New Jersey: Lawrence Erlbaum Associates.

Ross, S. M., & Morrison, G. R. (1989). In Search of a Happy Medium in Instructional Technology Research. *Educational Technology Research and Development*, 37 (1), 19-33.

Schimmel, B. J. (1988). Providing Meaningful Feedback in Courseware. In D. H. Jonassen (Ed.) *Instructional Designs for Microcomputer Courseware*, Hilsdale, New Jersey: Lawrence Erlbaum Associates, 183-196.

Scriven, M. (1993). *Hard-won lessons in program evaluation*, San Francisco, CA: Jossey-Bass.

Shneiderman, B., & Kearsley, G. (1989) *Hypertext hands on!* Reading, Massachusetts: Addison Wesley.

Southern Regional Educational Board (2002). *Criteria for evaluating Computer Courseware*, http://www.sret.sreb.org/criteria/courseware.asp.

Tessmer, M., Jonassen, D., & Caverly, D. A. (1989). *Nonprogrammers Guide to Designing for Microcomputers*, Englewood, Colorado: Libraries Unlimited Inc.