Search | Back Issues | Author Index | Title Index | Contents

ARTICLES

Usability of Hypermedia Educational e-Books

Paloma Díaz
Laboratorio DEI. Departamento de Informática
Universidad Carlos III de Madrid
Avda. de la Universidad 30, 28911 Leganés (Spain)
<pdp@inf.uc3m.es>

Abstract

To arrive at relevant and reliable conclusions concerning the usability of a hypermedia educational e-book, developers have to apply a well-defined evaluation procedure as well as a set of clear, concrete and measurable quality criteria. Evaluating an educational tool involves not only testing the user interface but also the didactic method, the instructional materials and the interaction mechanisms to prove whether or not they help users reach their goals for learning. This article presents a number of evaluation criteria for hypermedia educational e-books and describes how they are embedded into an evaluation procedure. This work is chiefly aimed at helping education developers evaluate their systems, as well as to provide them with guidance for addressing educational requirements during the design process.

1. Introduction

In recent years, more and more educational e-books are being created, whether by academics trying to keep pace with the advanced requirements of the virtual university or by publishers seeking to meet the increasing demand for educational resources that can be accessed anywhere and anytime, and that include multimedia information, hypertext links and powerful search and annotating mechanisms. To develop a useful educational e-book many things have to be considered, such as the reading patterns of users, accessibility for different types of users and computer platforms, copyright and legal issues, development of new business models and so on. Addressing usability is very important since e-books are interactive systems and, consequently, have to be designed with the needs of their users in mind. Evaluating usability involves analyzing whether systems are effective, efficient and secure for use; easy to learn and remember; and have a good utility [1].

Any interactive system, as e-books are, has to be assessed to determine if it is really usable as well as useful. Such an evaluation is not only concerned with assessing the user interface but is also aimed at analyzing whether the system can be used in an efficient way to meet the needs of its users [2] — who in the case of educational e-books are learners and teachers. Evaluation provides the opportunity to gather valuable information about design decisions. However, to be successful the evaluation has to be carefully planned and prepared so developers collect appropriate and reliable data from which to draw relevant conclusions.

This article presents a number of evaluation criteria for hypermedia educational e-books and describes how they are embedded into an evaluation procedure.

2. Evaluation Criteria for Educational e-Books

In this section, a number of evaluation criteria are described for hypermedia educational e-books that help educational developers in two ways:

- The criteria provide the basis for performing a formative evaluation to determine if the system is satisfactory or needs some changes.

- The criteria also provide guidance to software developers for addressing educational requirements (e. g., how to increase richness of didactic material).

Developers must assess the interface quality in order to detect usability problems or misunderstandings that need to be resolved to improve the interaction process. Moreover, evaluation must also focus on the e-book's utility, to analyze whether the e-book can be used in an efficient way to meet the needs of its users. Taking into account these premises, we propose a number of evaluation criteria for hypermedia educational systems [3] that can be applied to e-books. These criteria are based both on the works described in articles listed in the References section of this article [2, 4, 5, 6, 7, 8, 9] as well as based on the author's experience in the development of educational systems, such as CESAR [10], Now-Graduado [11] or CIPP [12].

2.1 Criteria to evaluate educational usefulness

Even more important than the quality of the user interface is the educational usefulness of a particular educational e-book. Developers of education materials must conduct tests to determine whether their e-books enable users to reach their learning and teaching goals. The following criteria can be applied to evaluate usefulness.

Richness. Garzotto, Mainetti and Paolini introduced the concept of richness [8], and the evaluation of educational e-books discussed in this article extends the concept in order to assess the richness of an e-book, taking into account parameters such as the following:

- *Information volume:* How many pages are there in the book, and how many information items are on each page?

- *Access richness:* Does the e-book offer users a variety of navigation paths and access tools? (For example, CIPP users, undergraduate students of Computing Engineering, who are learning the Pascal programming language are provided with five different browsing tools: 1) associative navigation through links embedded into the book pages, 2) sequential navigation following the hierarchical structure of the book, 3) a structural index based on this hierarchy, 4) a contextual index based on the structure of a Pascal program and 5) a search engine.)

- *Diversity of presentation and interaction styles:* Do the presentation and interaction styles meet the needs of different kinds of students with different learning styles? (For example, to learn sign language and written languages CESAR students, who are hearing-impaired children, are told stories using different media (video/image supporting sign language, images or text) depending on their learning competence.)

- *Kinds of exercises and interactive activities:* Does the e-book provide multiple kinds of exercises and activities to motivate students to take an active role in the learning process?

- *Scope:* How many educational objectives are fulfilled with the system?

Completeness. Measuring for completeness involves determining whether the system has an adequate number of content and interaction mechanisms to cope with the goals of different kinds of users. Some aspects to test when analyzing the system for completeness include:

- *Learning activities:* Have the designers provided learning tasks to fulfill the students' learning objectives? (Examples of useful learning activities include reading, creative writing, problem resolution and self-evaluation.)

- *Authoring support:* To what extent can the system be modified by users, both by students who edit, mark, annotate or restructure the didactic material, as well as teachers who require mechanisms to update the courseware or adapt it to their students' needs?

- *Communication support:* Will the system provide communication mechanisms among users? (Features that can be analyzed in this regard are the number of communication tools offered (e.g., e-mail, chats, videoconferencing) and the level of communication provided to the users (e.g., synchronous vs. asynchronous communication).)

- *Collaboration support:* Is there adequate support for collaborative activities? (Collaboration is not only about communication. Therefore other services must be provided, including definition and management of working groups as well as different user roles, each one with different responsibilities and abilities to access information.)

Motivation. It is also important to assess how students are motivated, not only to use the system but also to learn more about the subject being addressed. Aspects to take into account to improve the system with regard to motivating students are:

- *Self-evaluation mechanisms:* Are mechanisms available by which students themselves can assess their progress?

- *Exercises feedback:* Has the e-book been carefully designed to help students discover their errors and misunderstandings? (Feedback could be a message referring to the part of the course where the topic is explained or the assignment of a new exercise.)

- *Out-of-school activities:* Can the computer-based learning be related with other activities in the student's environment? (For instance, in Now-Graduado, at the end of each lesson, students — who are adults studying primary school subjects — are invited to perform non-computer-related tasks, such as visiting a museum or looking for a term in an encyclopedia. These activities strengthen the learning process and encourage students to continue their learning experience after they have finished a computer session.)

- *Adaptability:* Are adaptive mechanisms or personalized tools provided to make the system interface and function match the learning style of students?

Hypertext structure. This criterion is oriented towards analysis of structural properties as those proposed by Botafogo et al., Hatzimanikatis et al., and Yamada et al. [5, 6, 7]. It is fairly obvious that the node should be reachable or modular with regard to the hypertext. However, although there is no empirical evidence regarding their influence on system usability, other features like depth, imbalance, tree impurity and sequencing are also important and should be considered. Additional aspects of hypertext worthy of consideration include:

- *Connectivity:* Is the hypertext fully connected so that there is a hub node or page from which a path exists to reach all the pages of the e-book and vice versa? (The existence of isolated pages, which cannot be reached via links, reveals a negligent design.)

- *Modularity:* Is the resource structure modular in order to increase usability and ease maintenance?

- *Hierarchical structure:* Is the structure hierarchical? (A tree structure eases navigation and helps make the hypertext structure more explicit to the user.)

- *Balance:* Is the resource balanced? (A high degree of imbalance might suggest that some subjects are treated less deeply than others.)

Autonomy. With respect to multimedia components, Ficarra first defined autonomy as the degree of navigation freedom offered to the user [9]. Autonomy can also be redefined to include the degree of interaction freedom. Some aspects to analyze with regard to autonomy include:

- *Interaction freedom:* Does the e-book offer an appropriate level of autonomy for navigation and interaction? (For example, Now-Graduado, whose users had no previous computer knowledge or experience, includes circular links that connect distinct views of the same concept. When selected, circular links open the target node disabling all the node's links. Users can read the node and interact with it, but when finished, they come back to the source node.)

- *Help mechanisms:* Does the e-book provide clues and a help facility to assist users who may become disoriented? (For example, in CESAR, evaluators considered that associative navigation was not a problem as long as children could rely on a reference: the book structure.)

- *Autonomy degrees:* To what level is the resource autonomous with regard to different kinds of users? (For instance, associative navigation can be complemented with sequential access, as in CIPP.)

Competence. Competence is related to the ability to navigate through the system and to reach a particular goal. (See Ficarra [9].) Some aspects to consider concerning competence are:

- *Use levels:* Does the e-book offer different levels of use to meet different degrees of expertise? (Novice users require clues and assistance mechanisms that tend to bother expert users.)

- *Help mechanisms:* Does the resource provide interactive help mechanisms? (Such help has to be contextual so users receive the exact information they need at each moment.)

- *Adaptability:* Can the system adapt to the different interaction styles and are the contents appropriate for each type of user?

Flexibility. The ease with which the system can be used and maintained is evidence of its flexibility. Parameters for flexibility analysis include:

- *Accessibility:* What difficulties will users experience when trying to access the system? (Both the technical aspects (software and hardware platform required) as well as physical or personal aspects (e. g., disabilities, timing constraints, etc.) need to be analyzed.)

- *Modularity and structure of the system architecture:* Has the software system been structured in a modular and structured way so that future improvements will be easy and affordable? Moreover, can the system architecture be defined in such a way that modules can be reused for different applications? (Modularity considerably reduces the effort required to develop new courses.)

2.2 Criteria to evaluate user interface usability

As the communication channel through which the user comes into contact with the computer, the user interface has to make possible the performance of tasks by users. Criteria to evaluate the usability of hypermedia educational e-books are described below.

Aesthetic. How the inclusion of multimedia information is harmonized and used to enhance the comprehension of concepts is called its aesthetic. Analyzing the aesthetic takes into account parameters such as:

- *Legibility:* Are all the items included in each page presented in a way that takes into account the capabilities and limitations of different users and access platforms? (This parameter has to be considered both at the page and the content level.)

- *Rhythm of presentation:* Have the multimedia contents with an implicit duration (e.g. video or sound) been synchronized with the assimilation rhythm of users?

- *Density:* How many items does each page include? (Developers of educational resources need to include enough information per page without overloading users.)

- *Appropriateness:* Does the e-book have the appropriate look-and-feel for target users, taking into account the users' special characteristics (e.g., age, social and cultural background, disabilities) and needs? (For example, in Now-Graduado, where adult women were studying concepts usually taught to children between 5 and 14 years, graphical designers needed to be especially careful regarding the design of icons, images and the rest of the multimedia content to avoid a childish appearance that would have made mature users feel undervalued.)

Consistency. Consistency refers to the extent to which elements that are conceptually similar are treated equally by the application, while those that are different are treated differently. (See Garzotto et al. [8].) Consistent educational applications are easier to use and remember. Therefore, users can pay more attention to performing their tasks than to learning how to use the system. Consistency is analyzed with regard to the following:

- *Interface areas:* Is the visualization area organized into several areas (e.g., functions/buttons, contents and messages), and is each area always used for the same purpose?

- *Labels and messages:* Is the user interface, whether in written, pictorial or spoken form, consistent so that in the same situation, the same label or message is used?

- *Buttons, icons and menu items:* Are the design elements (buttons, icons, and menus) ordered following some criteria -- whether alphabetically or to minimize their semantic distance -- and are they also arranged in a unique bar or pop-up menu whose appearance remains stable, so the user always sees the same set of functions placed in the same positions? (Some adaptive applications that modify the icons bar or the menus according to some deductions made from the user's behavior tend to confuse and burden users.)

- *Interface clues:* Whenever an object attribute is used as an interface clue (e.g., color, background, sound), is it applied in a consistent way in all the pages where the object appears? (For example, underlining text in web-based applications can be considered a consistency error since underlined words are supposed to be links.)

Self-evidence. Self-evidence determines how easily users can guess the meaning and purpose of things with which they are presented. (See Garzotto et al. [8].) Self-evidence is mainly analyzed by checking to see how tangible the system structure and functions are. Some techniques that can be used to increase self-evidence include:

- *Metaphors:* Interface metaphors (e.g., the book, the story, the museum, the desktop) have been used in many educational systems to try to increase the e-book's tangibility.

- *Self-contained pages:* Pages should be designed so users can identify them easily.

- *Multimedia expressiveness:* Designers can take advantage of multimedia to make the pages of an e-book easier to recognise.

- *Links adequacy:* Links among pages must be coherent and useful, since an excessive number of unnecessary links makes it difficult to understand the e-book structure.

- *Meaningful naming of functions:* The actions users can perform should be identified by means of a name belonging to the application domain. Moreover, terminology used in texts, as well as in system commands and dialogues, has to be clear and unambiguous so the user can understand each message and concept.

Naturalness of metaphors. It is important to evaluate metaphors used in the e-book to see whether they improve communication with the user (see Ficarra [9]) or, conversely, whether they fail to convey all the features of the domain or communicate the features in different ways, both of which constrain and mislead users. For example, one of the most relevant conclusions of the evaluation of CESAR was that the use of book and story metaphors was a good choice, not only for usability purposes but for helping to socialize children (e.g., teaching them to share books or providing them with the opportunity of being told stories). Some aspects to be considered regarding the naturalness of metaphors are:

- *Conceptual appropriateness:* Does the metaphor encompass all the features of the application domain required to reach the user's goals? (If such features are not supported by the metaphor, it has to be improved.)

- *Effectiveness:* When a metaphor is used to represent certain features of the application domain, does it decrease system efficiency or fail to make the system more usable? (If so, then the metaphor needs to be relaxed. In this case, e-books should not assume the physical book metaphor in a rigid way, for instance imitating the appearance of a physical book, since the resulting e-book might be less efficient in terms of use. Moreover, it is necessary to consider whether or not users would benefit from the physical book metaphor.)

Predictability. The extent to which users can anticipate a system outcome may be thought of as its predictability (see Garzotto et al. [8]), that is, predictability is measured by the degree to which users know the kind of result they will get from a specific interaction. Predictability is different from self-evidence, since with self-evidence users can identify the purpose and function of each object they are presented yet they cannot be sure what will happen if they perform a particular action. The best way to increase predictability is to perform a task analysis with users in order to understand what they expect from each interaction and how results are supposed to be presented.

3. An Evaluation Framework for Hypermedia Learning e-Books

For evaluation criteria to be really useful, they should be integrated into a procedure that will guide developers during the assessment process. With this purpose, in this article the procedure defined in Catenazzi et al. [13] has been extended (see Figure 1) as discussed below.
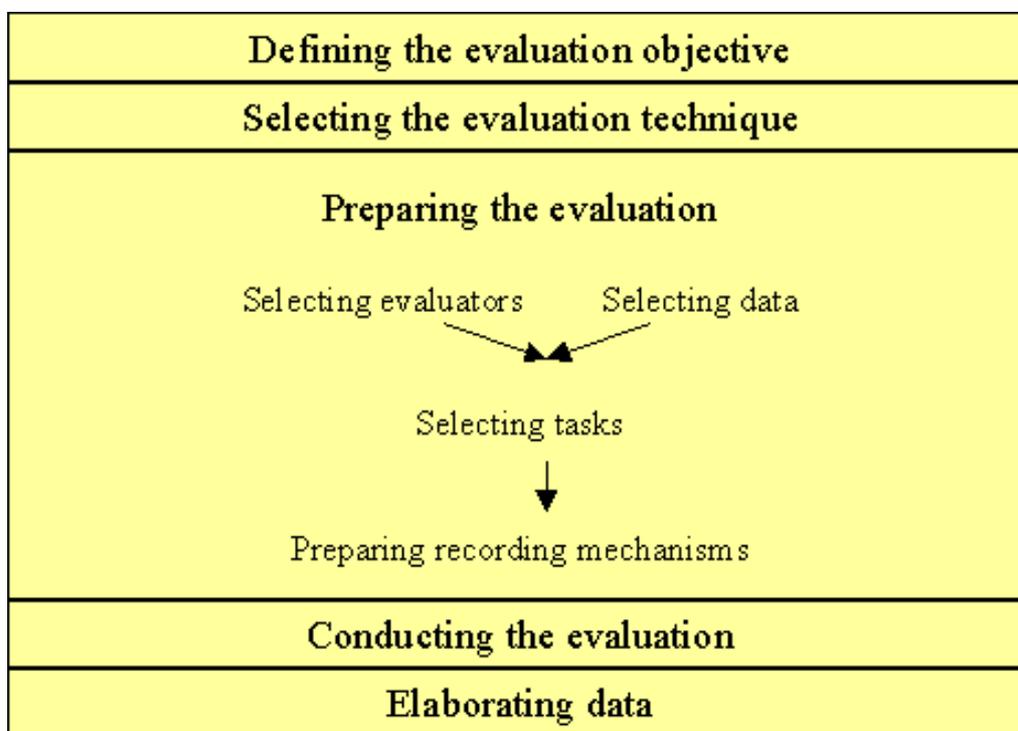
Figure 1. Evaluation procedure for e-books

Defining the evaluation objective. The evaluation of hypermedia educational e-books should be addressed in two different ways that emerge as two valid objectives:

1. Evaluating the educational usefulness

2. Testing the user interface

These objectives have to be put into more concrete terms in order to be measured, so designers should specify what they mean by educational usefulness and interface usability. That is, they need to determine the specific learning goals the e-book is expected to meet, as well as to assess the intended audience in terms of requirements such as learning styles, disabilities, background, age, software and hardware platform, and so on.

Selecting the evaluation technique. Several evaluation methods for interactive systems have been proposed in the literature, including analytic, expert, empirical and experimental procedures. The decision regarding which method to use depends primarily on such factors as what resources are available or what stage of development has been reached.

Preparing the evaluation. In this step, developers must do the following: decide what data to collect, select the evaluators, establish the tasks evaluators will carry out, and prepare mechanisms to record information on the evaluation process.

- *Selecting data.* To select the type of data that will be collected, the criteria proposed in Section 2 of this article can be used (establishing which criteria and which parameters for each criteria make sense depending on the evaluation goals). For instance, the usability criteria "naturalness of metaphors" would not be considered if no interface metaphors are adopted. Similarly, if the e-book does not consider learning as a collaborative task, the parameter "collaboration support" would not be necessary. Once the list of criteria and parameters is defined, some kind of priority or preference relationships can be determined among criteria and among parameters within each criteria to reflect their relevance to the e-book goals. Thus, it might be decided that to evaluate educational usefulness, "Motivation" is more important than "Autonomy", since one of the main goals of an educational tool is to encourage active student participation in the learning process. In turn, to assess "Motivation" it might be decided that "out-of-school activities" are more relevant than "exercises feedback", since they help students realize studying is an activity that goes beyond the limits of lectures.

- *Selecting evaluators.* Another step in preparing for the evaluation is to select the people who will take part in it. Developers should take into account the objectives, selected evaluation techniques and available resources. Evaluators with different backgrounds tend to provide richer information; thus, to evaluate an educational e-book requires not only users but also people working on educational technology and/or experts in user interfaces who can identify potential problems or deficiencies more easily than users can. For example, a student's opinion on the "completeness" criteria will be a priori less well-founded than the opinion of an expert in educational technologies.

- *Defining Tasks.* To ensure that evaluators will analyze all features and tools associated with the e-book, developers should establish a set of tasks to be performed. Different sets of tasks can be proposed for evaluators playing different roles. Moreover, it is necessary to test the tasks to be sure they make it possible to collect information on the criteria that is to be assessed.

- *Preparing recording mechanisms.* Finally, recording mechanisms are needed to collect data. Data can be obtained from each task execution using, for example, log mechanisms or interviews and questionnaires.

Conducting the evaluation. During this step, evaluation is carried out in one or more sessions, in a centralized or distributed manner, depending on objectives and available resources.

Elaborating data. The purpose of this step is to transform data collected, findings and recommendations for improvements to deliver a useful and usable e-book. The elaboration of data is usually based on statistical formulas, such as average, standard deviation, etc., for which conclusions related to the objectives of the evaluation are drawn.

4. Conclusions

Usability is a key concern when developing a hypermedia educational e-book. The e-book has to be analyzed in terms of educational usefulness and in the usability of the user interface. For this purpose, an evaluation process has to be carried out. Evaluation will provide data to derive relevant findings concerning the e-book's usability. This article describes an evaluation framework for hypermedia educational e-books that can be used to test the user interface for the e-book as well as its educational usefulness. The framework proposes a procedure whereby a number of criteria and parameters are used to assess the e-book (see Table 1).

Evaluation goal	Criteria	Parameters
Educational Usefulness	Richness	• Information volume • Access richness • Diversity of presentation and interaction styles • Kinds of exercises and interactive activities • Scope
	Completeness	• Learning activities • Authoring support • Communication support • Collaboration support
	Motivation	• Self-evaluation mechanisms • Exercises feedback • Out-of-school activities • Adaptivity
	Hypertext Structure	• Connectivity • Modularity • Hierarchical structure • Balance
	Autonomy	• Interaction freedom • Help mechanisms • Autonomy degrees
	Competence	• Use levels • Help mechanisms • Adaptivity
	Flexibility	• Accessibility • Modularity and structure of the architecture
User Interface Usability	Aesthetic	• Legibility • Rhythm of presentation • Density • Appropriateness
	Consistency	• Interface areas • Labels and messages • Buttons, icons and menu items • Interface clues
	Self-Evidence	• Metaphors • Self-contained pages • Multimedia expressiveness • Links adequacy • Meaningful naming
	Naturalness of metaphors	• Conceptual appropriateness • Effectiveness
	Predictability	• Task Predictability

Table 1. Evaluation Criteria for Hypermedia Educational E-Books

The set of criteria proposed in this article is also intended to help designers of other educational materials during the analysis and design stages. However, the proposed criteria should be viewed as an incomplete list, since evaluating utility and usability is a complicated and polyhedral undertaking, with many facets remaining open to discussion.

Acknowledgements

5. References

[1] Preece, J., Rogers, Y., Sharp, H. *Interaction Design: beyond human-computer interaction.* John Wiley & Sons, Inc. New York, 2000.

[2] Lee, S. H. *Usability testing for Developing Effective Interactive Multimedia Software: Concepts, Dimensions and Procedures.* Educational Technology & Society, 2(2), 1999.

[3] Díaz, P., Sicilia, M.A. and Aedo, I. Evaluation of Hypermedia Educational Systems: Criteria and Imperfect Measures. *Proc. of the International Conference on Computers in Education,* Auckland 3-6 December, 2002. 621-626.

[4] Mendes, M.E.X., Harrison, R. and Hall, W. Applying Metrics to the Evaluation of Educational Hypermedia Applications. *Journal of Universal Computer Science,* April 1998.

[5] Botafogo, R.A., Rivlin, E. and Shneiderman, B. Structural Analysis of Hypertexts: Identifying Hierarchies and Useful Metrics. *ACM Transactions on Information Systems,* 10(2), 1992, 142-180.

[6] Hatzimanikatis, A.E., Tsalidis, C.T. and Christodoulakis, D. Measuring the readability and maintainability of hyperdocuments. *Journal of Software Maintenance, Research and Practice,* 7, 1995. 77-90.

[7] Yamada, S., Hong, J. and Sugita, S. Development and Evaluation of Hypermedia for Museum Education: Validation of Metrics. *ACM Transactions on Computer-Human Interaction,* 2(4), 1995. 284-307.

[8] Garzotto, F., Mainetti, L. and Paolini, P. Hypermedia Design, Analysis and Evaluation Issues. *Communications of the ACM,* 38(8), 1995. 74-86.

[9] Ficarra, F.V.C. Evaluation of multimedia components. *Proceedings of the International Conference on Multimedia Computing and Systems,* Ottawa, 1997. 557-564.

[10] Díaz P., Aedo I., Torra N., Miranda P. and Martín M. (1998): Meeting the needs of teachers and students within the CESAR Training System. *British Journal of Educational Technology.* 29 (1), 1998. 35-46.

[11] Aedo, I., Díaz, P., Panetsos, F., Carmona, M., Ortega S. Huete E. A hypermedia Tool for Teaching Primary School Concepts to Adults. *IFIP WG 3.3 Working Conference Human Computer Interaction and Educational Tools.* Sozopol (Bulgaria). May 27-28, 1997. 180-188.

[12] Aedo, I., Díaz, P., Fernández, C., Muñoz, G. and Berlanga, A. Assessing the utility of an interactive electronic book for learning the Pascal programming language. *IEEE Transactions on Education,* 43(4), 2000. 403-413.

[13] Catenazzi, N., Aedo, I., Díaz, P. and Sommaruga, L. The evaluation of electronic books: Guidelines from two practical experiences. *Journal of Educational Multimedia and Hypermedia.* 6(1), 1997. 91-114.