Quality evaluation of educational multimedia systems

Binh Pham
Research Centre for Intelligent Tele-Imaging
University of Ballarat

Much effort has been spent on developing and producing multimedia systems, with insufficient attention to ensure their quality. This paper discusses different aspects of quality evaluation of multimedia systems, covering both cognitive and technical issues. The evaluation is viewed from three main perspectives: the product itself, how the product is used and the impacts of the product. Some relevant evaluation models for effective learning - in particular, objective-based, decision-based, value-based and naturalistic approaches - are examined and adapted for the purpose.

Introduction

The capability of integrating different types of data (text, images, sound, video) on more powerful computer platforms have encouraged a rush to develop multimedia systems. Currently, there are five main application areas: education and training, commercial services, entertainment and communications. The initial excitement and novelty seem to have waned after the availability of a plethora of mediocre multimedia systems on the market. Although much time, efforts and resources have been spent on the development of these systems, little attention has been focused on the overall evaluation of their quality. Many developers of multimedia systems come from technical sectors, hence they tend to place emphasis on improving technical aspects such as ways to store, retrieve, present material in an attractive way and to navigate within the systems. On the other hand, developers who are domain experts are more concerned with specific contents while not taking full advantage of what advanced technologies can offer. Criticisms of multimedia products have ranged from being boring, restricted to insubstantial and flashy. Generally, it is felt that these products are still too closely related to traditional source media and have not successfully broken a new ground to achieve innovative products.
Although the contents of multimedia systems vary with different domains, these systems share a general aim of communicating knowledge to users in more interesting and effective ways. Since multimedia technology is still relatively new, no adequate evaluation guidelines yet exist to ensure the quality of its products. On the other hands, much research effort has been devoted to the investigation of evaluation and assessment models for effective learning. In addition, educational multimedia systems (in the form of either coursewares or games) make up a large share in the multimedia market. It is therefore natural to ask if such evaluation models and assessment methods can be adapted for evaluating multimedia systems for educational purposes in particular, and to other application domains in general.

A good opportunity for exploring this problem arose when we received a CAUT grant for a project to design and develop a multimedia software package for teaching interactive 3D computer graphics (Pham et al., 1996). The aims of the software were to supplement and enhance the understanding and hands-on experience of students by conveying and demonstrating the fundamental concepts of computer graphics in a stimulating and intuitive way. The environment was also required to be learner-centred and to engage learners in an experiential and problem-based approach to learning. These aims appear to be typical of many educational multimedia software, which seek to provide sophisticated learning environments to empower the learner. The software may thus be viewed as a cognitive tool to overcome the limitations of existing environments, and to facilitate learners in the process of ‘doing’ (e.g. thinking, manipulating, constructing, testing, analysing, synthesising and reflecting). The guidelines for evaluation that we were seeking have two main purposes: to assist with the design, construction and improvement of the software; and to examine its value in a wider context.

In this paper, a global view of multimedia systems is examined and a systematic evaluation framework for educational multimedia systems is proposed, which covers both cognitive and technical aspects. The next section discusses an evaluation framework from three main perspectives: the product itself, how the product is used and the impacts of the product. It is then followed by the examination of some relevant evaluation models for effective learning - in particular, objective-based, decision-based, value-based and naturalistic approaches - with the intention to determine their suitability for the purpose of evaluation of multimedia systems.
**Evaluation framework**

Developing a multimedia system is a resource-intensive activity in terms of time, budget, technology and professional expertise. The potential richness of these products makes it difficult and costly to have an adequate evaluation of quality. Generally, testing is performed in various stages to ensure that certain prescribed criteria are fulfilled. The concept testing is carried out first by examining needs and markets, before a prototype is designed and developed. The feedbacks obtained by performing tests at various developmental stages assist with design improvements and further development. Useability, field trials and acceptance testing are then performed before the product is finally released for use. This kind of testing is essential for any software development and predominantly aims to ensure that the product functions well technically to serve an application specified by a customer. Quality evaluation, on the other hand, implies a wider context which concerns with the value or merit of the product, how good the content and structures of the product are, and how well-designed and appropriate tools are for achieving the objectives of the product. Thus, a set of criteria need to be defined to specify what qualities to be evaluated, the procedures to apply the evaluation and the way to interpret the results.

Reeves (1992) highlighted the importance of evaluation of multimedia within the context of its use and the characteristics of users. He also suggested that multimedia should be designed to support the principles that learning involves knowledge construction where new knowledge is built upon existing knowledge and within meaningful contexts. Thus, formative experimentation can be carried out (e.g. using interviews, observations and analysis) to examine whether these goals are attained. Although these observations are important, they do not form a comprehensive list of relevant aspects that should be explored. To achieve this, we propose to examine an evaluation framework from three perspectives: the product itself, how it is used and the impacts it exerts in a wider context.

**The product**

There are three major elements of an educational multimedia system that exert profound effects on its quality: knowledge content, the ways knowledge and tasks are represented and organised, and technical tools used for conveying and constructing knowledge. However, the development and evaluation of these elements cannot be taken in
isolation from each other, and especially, independently of the objectives of the system. In fact, the analysis of objectives should be the first task to be considered because they largely influence the choice of teaching or learning method, how material content and learning tasks are structured, how computer software is designed and how testing and evaluation are carried out during development and at the completion of the systems.

Objectives

There have been many attempts to classify types of learning and to construct learning models which present ways to explain and categorise learners’ behaviour and mental process. They may be viewed from four major perspectives: learning outcomes, cognitive, affective, and psychomotor. One of the most well-known work on the conditions of learning was by Gagne, which presented the view that learning outcomes can be systematically identified and measured in terms of knowledge by performing tasks and sub-tasks which are organised in a hierarchical fashion (Gagne & Briggs, 1979; Gagne, 1985). Based on this theory, Price (1991) introduced a sequence of prescriptive guidelines which have been used extensively for traditional instruction design. However, this approach was later challenged by the belief that the cognitive development process is cumulative and not necessary hierarchical (Hoffman, 1997). This view not only shifts the educational emphasis from ‘teaching’ to ‘learning’, but has also opened up the scope for structuring information and its access. It also permits designers to take better advantage of the non-linear access capability offered by multimedia technology.

The affective and psychomotor objectives deal with more subtle aspects of learners’ emotional and psychological responses which assist and stimulate commitment in learning. One conceptual framework which can be useful for analysing these objectives is Keller’s motivation theory (Keller & Kopp, 1987). His ARCS model which is based on four variables - attention, relevance, confidence and satisfaction - can also be used to facilitate the development of learning strategies and the technical design of the user interfaces and learners’ interactions with the system. This model encourages the integration of material and tasks within a familiar and real-life context and greater care in the selection of media and methods for presentation in order to make the learning experience more interesting and meaningful.
One concern is that some current educational multimedia systems may have placed too much emphasis on the affective and psychomotor aspects, at the expense of learning outcomes and cognitive development. The goals of knowledge acquisition might be diluted in the quest to enthuse and excite learners using spectacular effects provided by images, animation, video and sound. This could be a backlash reaction to the ‘dullness’ of traditional teaching style which relies largely on prescriptive instructions to produce pre-defined learning outcomes. A good educational multimedia system must not lose sight of educational objectives while taking advantage of what advanced technology can offer. The design of such a system for a specific subject domain should begin with the identification of knowledge to be acquired according to learners’ cognitive development process, before addressing how the activities or tasks to be structured, what material to be used and how to carry out these activities in order to satisfy the affective and psychomotor objectives. The choice of media and technical capabilities should be made according to this aim. Thus, one major concern is to evaluate how well different aspects of the product serve to achieve the objectives for knowledge acquisition.

Knowledge content and tasks

Once the objectives are identified, the next step is to determine the scope of material and tasks that are appropriate for conveying knowledge. Traditional curricula such as those provided in text books are often organised in a bottom-up hierarchical fashion, where knowledge is obtained by following a pre-defined sequence. Each chapter which provides knowledge on a concept, is composed of sections, where each section explains different ideas that are required to understand the concept. The sections may be independent of each other or built on each other sequentially. However, the concept can only be completely understood by going through all sections.

The concept of a spiral curriculum is based on the belief that knowledge is obtained progressively from general to specific and the information within a knowledge structure is inter-related. Thus, the content may be structured in stages, each of which is composed of self-contained modules presenting information concerning a complete concept at a different level of details. Each module introduces more complexity and competency by building on the previous module. The main advantage of this approach is that learners can still obtain some level of competency if they stop at any module or any stage. This contrasts with the traditional
hierarchical curriculum, in which each module does not attempt to give a complete view of a concept, hence learners might not have accumulated sufficient knowledge to be of much use to them.

The elaboration theory (Reigeluth, 1991) provides some strategies to organise the sequence of learning material to facilitate the assimilation of new knowledge. The underlying beliefs are that knowledge is more meaningful to learners if it is introduced using real-life problems and social context, and can be better understood if it is related to and built on existing knowledge. These strategies also emphasise active learning by encouraging learners to have more participation and exploration in activities. One advantage is that the associative linking of ideas underlying this theory can be implemented most effectively within a multimedia and hypermedia environment. In this type of environment, an overview of a concept is introduced first, together with hyperlinks to other parts which provide further details. Learners are at liberty to choose their own learning path and in any order that interest them.

Despite the advantages of the spiral curriculum, this approach might not be the best for certain situations. Some subject domains are not suitable or quite difficult to be presented or learned in this fashion. The most obvious examples are in technical areas such as mathematics where knowledge has to be constructed sequentially, and it is often not possible to have a complete view of a concept without sufficient details. Hence some order of learning has to be imposed upon learners, and the traditional curriculum is more appropriate. However, in such cases, some limited linking of associated ideas can still be used if certain details can be assumed without making the concept being learnt incomprehensible. Another criticism is that although it is beneficial to integrate new knowledge within real-life context and to make the learning task as entertaining as possible, this can go too far. The multimedia systems might run the risk of becoming entertainment products which can offer only shallow knowledge and furthermore, induce learners to be unable to think deeply or in abstract terms.

The following questions need to be considered for evaluation:

- Are the knowledge content and its organisation appropriate for achieving the specified objectives?
- Are they pitched at the right complexity level for the users that the system attempts to reach?
• Do the tasks that are designed to convey this knowledge stimulate and enhance users’ capacity for learning?

**Representations and structures**

The availability of the Internet and its browsers has opened up access to a vast amount of information for the masses. However, an important question to be addressed is whether such information access would lead to the acquisition of knowledge in a constructive and structured way to assist with learning? There are three current main approaches to structure information within educational multimedia systems: instructivist, constructivist and free access. These approaches vary in the degree of freedom that learners have in choosing information to access and the amount of guidance in this process provided by the systems.

Traditional instructional design principles provides a set of formal rules to assist with the development of instructions in three stages: need analysis, choice of material and methods, and evaluation. The needs analysis aims to define the learning objectives of the instructions and the tasks required to achieve these objectives. The material and methods are then selected based on a sequence of activities that help to attain the required outcomes from each learning process. The main criticism of traditional instructional design principles is that they follow an objectivist approach which assumes that knowledge is independent of instructions and learners. Thus, learners’ characteristics are not taken into account. The very precise and prescriptive nature of the approach stifles the learning process by not allowing learners to have initiatives to explore and be responsible to their own learning. Furthermore, the linear characteristic of traditional instructional design models make it harder for flexible design and rapid prototyping of multimedia systems. It should be noted, however, that the question of whether creativity is hindered by instructional design methodology, is still a subject of heated debates (e.g. Dick, 1995; Rowland, 1995). In addition, some contemporary instructional design models have attempted to integrate a more emphasis on a human-centred approach (e.g. Dick & Carey, 1996).

The constructivist approach which is learner-centred, is built on the psychological theory of learning and cognitive development. The underlying belief is that an instructional control strategy exerted by a learner is more effective for building up knowledge than a control strategy that is pre-defined in computer programs by system designers. In this strategy, students gain knowledge by constructing tasks and
exploring the systems themselves. It has been argued that this encourages students to be more confident, independent and responsible for their own actions. It also helps them to retain the enthusiasm and knowledge learned. Cunningham et al. (1993) proposed seven principles for constructivist design which aim to encourage students to actively participate in the learning process, to apply knowledge to realistic and relevant situations, and to collaborate with others. The emphasis is on providing tools for students to retrieve, record and analyse information, instead of providing specific instructions. Social interactions also play an important role in this model, where students work in groups with some support from the teacher. The non-linear characteristic of the process of knowledge acquisition matches better with the technical characteristics of multimedia systems, thus makes the constructivistic approach more attractive for multimedia system design than the instructivist approach.

While the instructivist approach provides students with explicit instructions, the constructivist approach only supply information from which students construct their own knowledge. Information may be provided directly for the knowledge construction task or structured with the aim to facilitate this process. Another alternative is to provide information implicitly through examples and relevant questions. As the objectives of the learning process is less precise and the knowledge is obtained in an unstructured fashion, it is difficult to assess what students have learnt, and how well a system has achieved its objectives. Thus, there is an obvious need to devise new approaches for assessment that reflects students' interests and needs, in addition to current approaches that are instructor-oriented. Another issue to be considered is a perceived danger that the constructivist approach encourages extreme individualism, where students might create their own realities and meanings of the world, based largely on their own experiences (Duff & Jonassen, 1991; Perkins, 1991). These drawbacks are amplified in a free access approach similar to that provided by the browsers for the Internet. For this reason, free access should only be used sparingly and judiciously within educational multimedia systems.

It seems rather extreme to limit the task structuring of a system exclusively to either the instructivist or constructivist approach, and to deny the benefits of the other approach. A more optimal approach would be to use a mixture of variants of the constructivist approach or the instructivist approach, depending on the level of complexity of concepts to be learned and the ability of learners. The tasks can also be
structured in a progressive way, giving information to learners implicitly first, then gradually increasing the level of preciseness until explicit instructions are given, to cater for diverse abilities. Learners can thus switch between alternatives according to their own needs. The evaluation therefore would be best to focus on how these two approaches are balanced to achieve the needs of users with a wide range of abilities and backgrounds.

**Technical aspects**

The usefulness and attractiveness of a multimedia system depend not only on its content but how well technology is deployed to provide capabilities not available in traditional media. A slow, badly designed or unreliable system would quickly lose support of users. Technical evaluation needs to cover four main areas: speed, capacity, reliability and extensibility. Questions to be addressed include:

- How responsive is the system for interactive use?
- Is it fast enough to provide real-time response, and if not, is there any message to inform users of what is going on?
- What technical features are offered by the system?
- How reliable are they under extensive use by different types of users?
- Are these features better than those provided by other similar systems?
- Do they provide new functionalities or innovative way to perform a specific task?
- Were these capabilities be implemented in such a way that can be extended easily?

**Use of product**

The effectiveness of an educational multimedia system cannot be judged from the content of the product alone. How the product is used is of paramount importance. This is influenced mainly by the ways human computer interfaces and internal linkages between information structures are designed.

**Human computer interfaces**

Human computer interface issues are concerned with the choice of media, software and hardware tools and environment to enable the
product to do what users want in a pleasant and efficient manner. How to design good interfaces and how to evaluate and ensure their quality have been extensively covered in the literature, hence not dealt with in detail here (e.g. Nielsen, 1993; Shackel, 1986; Sutcliffe, 1995). In brief, proper use of many types of media may encourage attention, interest and commitment. It may also add imagination and foster creativity and aesthetic appreciation. However, injudicious use of this technology may backfire, reducing the effectiveness of the product or even turning users away. The evaluation may be carried out for a number of purposes: diagnostic (to identify poorly design features), comparative analysis (to test the effects of features or to compare different designs), or benchmarking (to test against certain recommended standards). The evaluation is normally expressed in terms of the degree of utility and useability, where utility examines the functionality and requirements of the system, and useability the requirements for human performance. The following questions need to be posed:

• How well can users carry out the targeted application using the interface?
• What is the degree of users’ satisfaction?
• How easy for users to learn ways to operate the system and remember them?
• How effective and efficient users can perform each of these specific tasks?

**Navigation and links to information structures**

Once the knowledge content and tasks have been determined, the choice of how to divide the required information into nodes, what kind of links to be used, and how much related information to be provided depend very much on the intended users and the types of applications. Some developers take an easy way out by mapping the material in the same manner as the way the source material is presented (e.g. presenting material hierarchically as it appears in a text book). This approach is very limited and unproductive because little is gained by just transferring the material from paper to electronic medium. Innovation can only result if the rich capability that technology is exploited, to create a more distinct approach to communicate the same material.
There are a number of models for structuring information in a multimedia or hypermedia environment from a technical point of view. The Amsterdam Hypermedia Model (Hardman et al., 1994) extends the Dexter Hypertext Reference Model (Halasz & Schwartz, 1994) which models a hypertext system as basic components with appropriate link structure. This model represents the system as three layers: within-component, storage and runtime. The within-component layer contains material and media belonging to a specific application, while the storage layer contains file structure and database organisation of components. Information in the within-component layer can be accessed via anchors which provide identifiers of different parts of material or media. The runtime layer deals with links, anchors and components. Links may be of three types: linear, hierarchical or network. In addition to structural links, related information may be accessed via associated links. The Amsterdam model extends the Dexter storage layer by adding multimedia data types and temporal data. It also provides ways to manage the synchronisation of different types of data. If links are achieved through concepts and not by specifying the location of anchor points and destination nodes, then much less effort will be required to construct links between associated information. However, this capability is currently available only for text data due to the difficulty of extracting information content from other types of data.

To communicate an application effectively, a few pertinent questions need to be addressed:

- How easy can users obtain knowledge or perform tasks by following the links provided by the system?
- Does information content provided in each node and its associated nodes facilitate relational understanding of concepts?
- How do such links and navigation methods provide more effective ways to disseminate knowledge than traditional media?
- Do they stimulate creative ideas and commitment?

**Impacts of product**

Generally, evaluation carried out for computer-assisted learning packages have been mainly concerned with three aspects: to determine how well they can be deployed within the curriculum, whether they have achieved intended objectives and whether they have performed
well in comparison with other similar products. These issues are worthwhile and certainly need to be examined for educational multimedia systems. However, the impacts may be far reaching for multimedia systems due to their potential richness. It is thus essential to evaluate the degree of innovation of a system as well as unintended outcomes that may result as spin-offs. For example, new ways for accessing information may change the way people think and work, or enable them to see certain aspects that had not been possible otherwise. Judging from the fact that the impacts of a computerised society have not been fully appreciated although computers have been around for more than three decades, we need to keep our mind open and not be restricted by what was intended before a multimedia system was developed.

Evaluation approaches

Numerous educational paradigms and evaluation models for instructional technology are available in the literature (for example, brief reviews and classifications may be found in Alexander & Hedberg, 1994; Reigeluth, 1987). One difficulty that has deterred multimedia system developers from adopting these models is the confusion they cause. Some models are not sufficiently explicit to allow easy implementation in practice. Others even appear contradictory and are not substantiated by validation studies. In addition, some models overlap significantly because although different words are used, they describe very similar concepts.

Alexander and Hedberg (1994) examined current practice in evaluation for various technology-based learning projects reported at two recent conferences in Australia and USA, and gave a brief review of four main approaches for evaluation: objective-based, decision-based, value-based and naturalistic. The objective-based approach determines the degree of achievement of educational objectives, without taking into account students' characteristics or abilities. The main intention is to obtain a precise summative evaluation to assist with further improvement of the educational activities such as curriculum planning or teaching methodology. The decision-based approach which allows contextual examination in detail, focuses more on the intermediate stages of product development by evaluating the goodness of decisions that are made at each stage. The value-based approach is more concerned with the overall merit and worth of the product, thus offers a broader perspective which may include major consequences not intended by the
designers of the product. All these three approaches concentrate mainly on the product and its outcomes, and do not deal with the effects on users explicitly. The naturalistic approach, on the other hand, is concerned with users’ views, interests and experiences. Information in this case is generally obtained through observations and interviews of users verbally or through questionnaires.

As these four evaluation approaches produce different insights into different types of learning outcomes or impacts, it is more advantageous to use them judiciously in combination in order to extract information more fully. Data collected for evaluation purposes may be of either quantitative (e.g. test scores and statistical measures) or qualitative type (e.g. symbolic representations in natural languages or verbose comments via observations or interviews of users). It has been argued that traditional quantitative data analysis based on objective measures such as learning outcomes is not only too limited but tends to categorise people into fixed types with different abilities. Hager and Butler (1996) suggested that a more effective assessment model should focus on the categorisation of tasks and on the cognitive development of users and their competency within a wider context than just the achievement of a specific task. This context may include areas such as skills in problem-solving, application, analysis, synthesis and in cooperation with other people. The belief underlying this model is that individuals have capacity to develop and improve their performance through learning. A purely quantitative model therefore would be insufficient for assessing such purposes.

**Conclusion**

Relevant issues concerning quality evaluation of educational multimedia systems from both technical and cognitive perspectives have been presented, with the aim to provide some concrete guidelines to assist with the development of these systems. It is hoped that deeper understanding of these issues in combination of ever-increasing capabilities of multimedia technology would inevitably result in future products which are more innovative and imaginative, and may exert deeper impacts on learning strategies and effectiveness.
References


Binh Pham
Research Centre for Intelligent Tele-Imaging
University of Ballarat
PO Box 663, Ballarat VIC 3350 Australia