Multimedia Evaluations Based on Cognitive Science Findings

Eshaa M. Alkhalifa
University of Bahrain, Bahrain

INTRODUCTION

Multi-media systems waltzed into the lives of students and educators without allowing for the time required for the development of suitable evaluation techniques. Although everyone in the field is aware that judging this type of teaching software can only come through evaluations, the work done in this regard is scarce and ill organized. Unfortunately, in many of the cases the evaluation forms were just filled in by instructors who pretended they were students when they went through the tutorial systems (Reiser et al., 1994).

BACKGROUND

In the early days, some researchers regarded the evaluation of the program's functional abilities and efficiency to be important, so they defined them as formative evaluation, and they also defined the effectiveness of the system as summative evaluation (Bloom et al., 1971; Scriven, 1967).

Others believe the evaluation of the system is unimportant, so they focused on the latter by comparing student performance in pre- and post-test questionnaires prior to and following the use of the system, learning style questionnaires that targeted their learning preferences and a subjective questionnaire that investigated whether students like the system (Kinshuk et al., 2000). Unfortunately, many of the pre- and post-tests resulted in no significant1 differences in student grades when multimedia is compared to classroom lectures or to carefully organized, well-illustrated textbooks (Pane et al., 1996). These disappointing results caused researchers to question whether or not the correct evaluation questions are being asked; for example should the test be of interactivity versus lack of interactivity, or should one compare animation with textual media (McKenna, 1995)? If Pane et al. (1996) were aware of the work done by Freyd (1987), who studied the cognitive effects of exposing subjects to a series of still images to find that they are equivalent in the reactions they elicit to being exposed to a moving picture, then perhaps they would not have asked whether animation is equivalent to a textbook with carefully set images of all stages.

Since the problem that arose is the evaluation question, researchers continued to alter it in order to recognize what should be emphasized. Tam et al. (1997) proposed a three-part evaluation procedure that includes peer review, student evaluation as well as pre- and post-testing (Tam et. al., 1997). They were not able to get rid of the pre- and post-test evaluation, as it is the primary test for how much learning was achieved, and they still got no significant differences.

At this stage, researchers recognized that evaluations did not target the appropriate level of detail, so Song et al. (2000, 2001) presented empirical support that animation helps reduce the cognitive load on the learner. They also showed that multi-media is more effective in teaching processes than in teaching conceptual definitions, while textual presentations are better at the latter. However, all this was done in very limited test domains that lacked the realistic world of an educational system. Albalooshi and Alkhalifa (2002) implemented some of these ideas in addition to offering both textual representations and animations within the same screen to students. This supports individual learning preferences while offering multi-media systems as a cognitive tool. Such a tool requires an evaluation framework that is well informed of the justification behind its design and the way its main modules interact.

A 3-DIMENTIONAL FRAMEWORK FOR EVALUATION

In the reported cases, most of the evaluated systems failed to reflect their true abilities because some aspects of the design or effects were neglected. Consequently, a complete framework of evaluation is required to take into account all issues concerning the software and the learning process. Evaluation questions can be channeled into three main dimensions of evaluation that could then be subdivided into the various methods that form possible criteria that guide the evaluation process.

Copyright © 2005, Idea Group Inc., distributing in print or electronic forms without written permission of IGI is prohibited.

Table 1. A three-dimensional framework of evaluation

1st Dimension: System Architecture

This dimension is concerned with the system's main modules, their programming complexity as well as their interactions. Evaluation within this dimension should be performed in any or all of the following methods:

- Full description of system modules and complete check of interaction.
- Expert survey of the system filled by experts or educators.
- Student evaluations to consider their perspective of the system.
- Architectural design must be based on cognitive science findings rather than chance.
- Everything else concerning the system design such as cost analysis and portability.

2nd Dimension: Educational Impact

This dimension is concerned with assessing the benefits that could be gained by students when they use the system. Classically, these are done in pre- and post-tests, and this is carried on in this framework with more attention given to detail.

- Students grouped according to their mean grade in a quiz.
- Post-tests are used to compare one group with system only and another classroom only. A third group attends the classroom lecture with the class group and does a pre-test, and then uses the system before doing a post-test for comparison with the other two.
- Questions in the pre/post-tests must be mapped to each other to test the same types of knowledge, mainly consisting of declarative and procedural knowledge.
- The tests should best be attempted with students who were never exposed to this material previously to assess their learning rate.

3rd Dimension: Affective Measures

This dimension is mainly concerned with student opinions on the user friendliness of the system and allows them to express any shortcomings in the system. This could best be done through a survey where students are allowed to add any comments they wish freely and without restraints.

The framework will be explained through a case study that was performed of a data structure tutorial system (DAST) that was developed and evaluated at the University of Bahrain (AlBalooshi & Alkhalifa, 2003). The process started with a pre-evaluation stage, where students were all given a test and then were divided into groups of equivalent mean grades. This was done to allow each group to have members of all learning levels.

Then the pre- and post-tests were written to ensure that one set of questions mapped onto the next by altering their order while ensuring they include declarative questions that require verbalization of how students understand concepts as well as procedural questions that test if students understand how the concepts can be applied. Last but not least, a questionnaire was prepared to allow students to highlight what they regard as any weak areas or strong areas based upon their interaction with the system. The evaluation procedure for students is shown in Figure 1. Educators were also asked to fill in an evaluation form as experts.

ANALYSIS OF RESULTS

First of all, student grades were analyzed using the Analysis of Variance (ANOVA) test. This test allows the evaluation of the difference between the means by placing all the data into one number, which is F, and returning as a result one p for the null hypothesis. It will also compare the variability that is observed between conditions to the variability observed within each condition.

The statistic F is obtained as a ratio of two estimates of students' variances. If the ratio is sufficiently larger than 1, then the observed differences among the obtained means are described as being statistically significant. The term "null hypothesis" represents an investigation done between samples of the groups with the assumption that additional learning will not occur as a result of the treatment. In order to conduct a significance test, it is necessary to know the sampling distribution of F given that the significance level needed to investigate the null hypoth-

3 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/multimedia-evaluations-based-cognitive-science/14560

Related Content

Internet: A Right to Use and Access Information, or a Utopia?
Inban Naicker (2008). *Information Communication Technologies: Concepts, Methodologies, Tools, and Applications (pp. 1306-1327).*
www.irma-international.org/chapter/internet-right-use-access-information/22740

The Impact of Information and Communication Technologies on Economic Growth and Electricity Consumption: Evidence from Selected Balkan and Eastern European Countries
Burcu Berke, Gülsüm Akarsuand Gökhan Obay (2017). *Information and Communication Overload in the Digital Age (pp. 176-200).*
www.irma-international.org/chapter/the-impact-of-information-and-communication-technologies-on-economic-growth-and-electricity-consumption/176570

Exploring Social Media for Health and Wellness: A Health Plan Case Study
Sinjini Mitraand Rema Padman (2012). *Journal of Cases on Information Technology (pp. 42-64).*
www.irma-international.org/article/exploring-social-media-health-wellness/71812

Digital Literacy and the Position of the End-User
Steven Utsiand Joost Lowyck (2005). *Encyclopedia of Information Science and Technology, First Edition (pp. 875-879).*
www.irma-international.org/chapter/digital-literacy-position-end-user/14352

Actor-Network Theory and Adoption of E-Commerce in SMEs
Arthur Tatnalland Stephen Burgess (2005). *Encyclopedia of Information Science and Technology, First Edition (pp. 38-41).*
www.irma-international.org/chapter/actor-network-theory-adoption-commerce/14207