Evaluating Educational Software Environments

Petri Gerdt
pgerdt@cs.joensuu.fi

Ramin Miraftabi
ramin@cs.joensuu.fi

Markku Tukiainen
mtuki@cs.joensuu.fi

Department of Computer Science
University of Joensuu
PL 111, 80101 Joensuu, Finland

Abstract
We present an outline of a checklist, the TUP-model, that covers the technological, usability, and pedagogical aspects of the educational environment so that more complete evaluations may be created. We base our model on our experience in usability evaluations and work with teachers in evaluating educational environments.

1. Introduction

Teachers selecting software environments for use in their own educational settings face a daunting task with the ever increasing number of educational environments available. By educational (software) environments we mean any software system or other electronic material that is used in an educational setting. Trying to find reviews of the plethora of environments available is even more challenging. Comparing the systems using the found reviews is currently even more difficult as there are no existing widely used evaluation methods.

2. Existing checklists

A number of generic checklists have been created for the evaluation of pedagogical and usability aspects or technical issues of software environments. In most cases generic checklists are adequate enough as creating domain specific checklists is a time-consuming and expensive task which is often left undone. Of the various checklists that have been created to help in the evaluation of educational environments we will take a brief look at the Delta checklist [7] and the constructivist checklist [2].

The constructivist checklist is directly tied to the underlying theory and focuses almost exclusively on the pedagogical aspects of the system being evaluated. The Delta checklist takes a slightly broader perspective with a strong emphasis on evaluating multimedia and visual appearance.

Ravden and Johnson [5] have created a method for evaluating the usability of a computer application. Our experiences with a slightly modified checklist [8] enable us to argue that the checklist is usable with end-users.

We have also used the Delta and modified Ravden and Johnson checklists with teachers creating the evaluations. The experiences obtained from these efforts have prompted us to begin the development of a new evaluation model.

3. The TUP-model

The experiences we have of using various checklists in the evaluation of educational environments has shown that they all concentrate mainly on one aspect, mainly the pedagogical aspect. Thus we propose to create a checklist that takes into consideration the instructional, technical, and usability aspect of educational environments. Some earlier work in this field exists (e.g. [6] and [3]). Of these the work done by Squires [6] is a step closer to our approach in which he extends heuristic checklists to include the evaluation of the environment’s suitability to a given learning situation.

Our proposed checklist concentrates equally on the three aspects, thus the TUP-model. The three aspects are strongly interrelated, as shown in Fig. 1. As can be seen, both usability and pedagogy build upon the foundation of the technological aspects. Usability and pedagogy are interdependent and all three aspects have the functionality of the environment at their core.

It is important to note is that the whole evaluation is based on a single scenario of use. While use scenario’s description belongs to the pedagogical aspect, we feel that it must be kept in mind during the whole evaluation. This means that undue generalizations must be avoided and the answers should relate to the use scenario.

For each of the three aspects we need to consider several different issues. In the technical aspects we must consider the availability, maintainability, and ease of initial setup and administration of the environment. In addition, the environments security and respect for the privacy of its users must
be evaluated, especially if the environment tests users and stores the scores. An additional important technical issue that needs consideration is how well the environment supports material sharing and reuse with other parties.

The usability aspects include evaluating the learnability, efficiency, and memorability of the user interface. These are three of the five main focus areas of usability according to [4]. The two missing focus areas are subjective satisfaction and the error rate. We feel that these two areas can not be evaluated in an educational environment as dissatisfaction with a learning result may lead to a poor evaluation of the usability and vice versa. The error rate may also be increased by problems related to learning the subject matter and not the interface.

In addition to the three focus areas we need to evaluate how well the environment takes into account perceptual and motor factors as well as the information factors related to the environment’s use. An overlap between the technical and usability aspects can be found when the visual aesthetics and internationalization abilities of the environment are taken into account.

The basis of the evaluation of the pedagogical aspects in TUP-model is the use scenario, which outlines the situation and context in which the evaluation of the software is done. In this context several issues about the software environment are evaluated: supported educational approach(es), pedagogical soundness of content, supported types of interaction [1]. Focus on how the environment motivates its users (e.g. internal vs. external motivation) and the environment’s support for self-evaluation chart more learner-related characteristics is needed. Another important issue is the possible integration of the evaluated software environment with other (educational) materials.

4. Discussion

We have presented an outline of a new evaluation checklist that conforms to the TUP-model, also presented in this paper. It is based on the equal weights of technical, usability, and pedagogical aspects.

In order to support our evaluation checklist we are developing it simultaneously with an on-line environment that allows users to evaluate educational environments. A centralized repository of evaluations will help educators in locating evaluations made by their peers. As we have already noted, evaluations made by peers in similar situations are more useful to teachers than those of experts. We also feel that peer evaluation is the best policy available. While supporting the possibility of several reviews for each evaluated environment, we are also considering various approaches that would enable ranking of reviews.

Many of the Websites that enable reviewing products (such as Amazon.com) also enable reviews of reviews. Enabling reviews of the evaluations will help the users in locating evaluations that have been useful for others. We are also considering how a Slashdot (slashdot.org) style ranking of evaluators could be implemented. In Slashdot’s ranking system community members can rank pieces written by different authors and these ranking affect the authors overall karma and thus the default rank that all texts written by the author. This ranking would help in establishing trust between users who will quite possibly never meet or even hear about another in other environments.

References