International Journal of Business, Human and Social Sciences
ISSN: 2517-9411
Vol:1, No:6, 2007

Augmented Reality in Schools: Preliminary Evaluation Results from a Summer School

Alexandru Balog, Costin Pribeanu, and Dragos Iordache

***Abstract*—**Formative usability evaluation aims at finding usability problems during the development process. The earlier these problems are identified, the less expensive to fix they are. This paper presents some preliminary results from a formative usability testing of the 1st prototype developed for the ARiSE (Augmented Reality in School Environments) project.

***Keywords*—**AR-based educational systems, formative evaluation, usability evaluation, user testing.

I. Introduction

THE implementation of innovative pedagogical/educational practices is a response to the social needs for educational change [3]. Traditional methods of educating students have well-proven advantages, but some deficiencies have also been detected. A typical problem has been how to engage students with appropriate information and communication technologies (ICT) during the learning process [6].

Augmented Reality is a variation of Virtual Reality (VE) that supplements reality, rather than completely replacing it. According to Azuma [2], AR systems are featuring an integration of real and virtual (computer generated images) into real environments, real time 3D interaction and targeting all senses (visual, auditory and haptical).

AR technologies are expensive and require a lot of research and design effort to develop visualization and rendering software. On another hand, the mix of real and virtual requires appropriate interaction techniques that have to be tested with users in order to avoid usability problems.

Formative usability testing is performed in an iterative development cycle and aims at finding and fixing usability problems as early as possible by testing the software with a relatively small number of users. The earlier these problems are identified, the less expensive development effort to fix them is required. As Hix et al. [5] pointed out; this kind of user-based statistical evaluation can be especially effective to support the development of novel systems as they are targeted at a specific part of the user interface design.

ARiSE is a research project that aims at creating an augmented reality technology for schools by adapting a virtual showcase used in museums. ARiSE will develop interaction scenarios for learning and associated software prototypes in order to assess the pedagogical effectiveness of the AR technology. The project is carried on in a consortium of seven partners: Franunhofer IAIS (Germany) – coordinator, Siauliay University (Lithuania), AccrossLimited (Malta), ICI Bucureşti (Romania), Czech Technical University in Prague (Czech Republic), Siauliai City Juventa Basic-School (Lithuania) and Rabanus-Maurus Gymnasiums Mainz (Germany).

The 1st ARiSE prototype has been tested with users during a summer school organized in Hamrun, Malta. All partners participated to the summer school. The objectives of the test were to assess the pedagogical effectiveness and usability of the prototype. Several evaluation techniques have bee used: observation, usability questionnaire and focus group.

This paper presents the usability evaluation results from the ARiSE summer school. The results are based on the usability questionnaire filled in by students. The rest of this paper is organized as follows. Related work in this area is presented in the next section. The evaluation context (platform, participants and tasks) is briefly described in section III. The evaluation results are presented in section IV. The paper ends with conclusion in section V.

II. Related Work

In the literature, there is no usability method yet specifically designed for Augmented Reality systems for school environments, but there are some researches focused on the evaluation of e-learning systems.

Thereby, Ardito et al. [2, 3] proposed a systematic evaluation method of e-learning web-based systems usability – SUE (Systematic Usability Evaluation), which combines the usability inspection with users-testing. The main novelty of this methodology is the use of evaluation patterns, called abstract tasks (AT). The evaluation patterns describe how to estimate the compliance of application components with a set of attributes and recommendations, which are initial identified for a certain application class. ATs guide the inspector's activities describing which objects of the application to look for and which actions to perform during the inspection in

Manuscript received July 31, 2007. This work was supported in part by the FP6-027039 research project (ARiSE) and the FP6-507609 european research project (SIMILAR).

Alexandru Balog is with the National Institute for Research and Development in Informatics – ICI Bucureşti, Bd. Mareşal Averescu No. 8-10, 011455 Bucharest, Romania (phone: +40-(0)21 3160736; +40-(0)21 3161030; e-mail: alexb@ici.ro).

Costin Pribeanu is with the National Institute for Research and Development in Informatics – ICI Bucureşti, Bd. Mareşal Averescu No. 8-10, 011455 Bucharest, Romania (phone: +40-(0)21 3160736; +40-(0)21 3161030; e-mail: pribeanu@ici.ro).

Dragoş Daniel Iordache is with the National Institute for Research and Development in Informatics – ICI Bucureşti, Bd. Mareşal Averescu No. 8-10, 011455 Bucharest, Romania (phone: +40-(0)21 3160736; +40-(0)21 3161030; e-mail: iordache@ici.ro).

International Journal of Business, Human and Social Sciences
ISSN: 2517-9411
Vol:1, No:6, 2007

order to analyze such objects.

The main advantages of SUE methodology is that it combines usability inspection with user-testing, which is both a fast and efficient solution (discount usability evaluation). The fact that the user-testing proportion is relatively low could be considered an inconvenience. Another disadvantage is represented by the controvertible effectiveness of evaluation made with beginner evaluators, based on ATs. Thereby, this method is effective for new module / content pages, on a platform which was already evaluated.

Amershi et al [1] mentioned the following pedagogical goals which support the designing of an interactive environment of visualization for computer assisted learning.

- Understanding of the target area by student.
- Support for different learning skills, learning styles and levels of knowledge.
- Motivation and increasing the interest for the teaching topic.
- Promotion of an active engagement in using interaction tools.
- Supporting different learning scenarios, including demonstrations in classroom, homework and exploration.

The authors propose the following usability objectives:

- Easy to learn, to understand and to use.
- Easy to integrate the lesson into a course.

According to Jacob Nielsen [11] an artifact is fruitful if is both useful and usable. Starting from this definition Sillius and Tervakari [13] propose the next attributes for an e-learning system utility:

value added: organization of teaching process, development of quality of teaching, development of learners' skills, testing and development of educational ICT.

pedagogical usability: support for organization of the teaching and studying, support for learning and tutoring processes as well as achievement of learning objectives and support for the development of learning skills (interaction with other actors, growth of learners' autonomy and self-direction).

Petri Nokelainen [12] makes a distinction between technical and pedagogical usability in the context of usability evaluation of digital learning material. According to Nokelainen "when evaluating the usability of a specific *learning platform* (learning management system), we can evaluate how easy it is to use the platform itself (technical usability), or what kind of learning material it enables the users to produce (pedagogical usability). When evaluating the usability of a *learning unit* or a *learning object*, we assume that each learning unit has its own interface relating to the content, and learning content based on a certain learning goal. When evaluating the pedagogical usability of a learning unit, we must try to control the effect of the pedagogical solutions of a learning platform."

Another system is „Challenges", a paper-based tool developed by Agnes Kukulska-Hulme and Lesley Shield [9] which aims to supporting content experts in developing their understanding of typical pedagogical usability problems and possible solutions. The tool comprises two major sections. The first introduces users to key usability terms and concepts such as the lifecycle of the course website, from design to maintenance via implementation. It also offers practical 'hints and tips' about planning and design issues like the importance of adequate planning before handing the website over for programming, and effective methods of testing ideas – for example, paper prototyping – before committing the course developers to a specific approach. The second section of "Challenges" takes 10 key usability issues or *Challenges*, offering comprehensive information and advice about each.

III. The Evaluation Context

A. The AR platform

The AR platform consists of 4 independent modules organized around a table on which real objects are placed. The platform has been registered by Fraunhofer IAIS under the trade mark Spinnstube®.

In Fig. 1, the photo of a module is presented.

Fig. 1 A module of the ARiSE platform

The project will implement three prototypes based on three interaction scenarios. The 1st prototype is targeting Biology. The real object is a flat torso of the human digestive system. A paddle has been used as interaction tool that serves for three types of interaction:

- Pointing on a real object.
- Selection of a virtual object.

- Selection of a menu item.

B. Participants and Tasks

The data model is capturing representations of domain. The test has been conducted at the office of the partner AcrossLimits in Hamrun, Malta.

Five user teams from 4 countries (Germany, Lithuania, Malta, and Romania) participated at the summer school with a total of 20 students from which 10 boys and 10 girls. None of the students was familiar with the AR technology. 16 students were from 7th class (13-14 years old) and 4 from 11th class (16-17 years old).

The participants have been assigned 4 tasks: a demo program and three exercises. The tasks have been presented via a vocal user interface in the national language of students. According to the test plan, each team should test the prototype in two working sessions: demo + 1st exercise and 2nd + 3rd exercise.

IV. Results

A. Analysis of Usability Questionnaire Data

The usability questionnaire consisted of 12 scale items. The first 10 items were related to three dimensions of AR platform quality (e.g. interaction, usability, and capability). The remainder two items were to assess students' perceived overall easy of use and exciting, respectively. Respondents were required to rate the questions on a 5-point Likert scale (1-strongly disagree and 5-strongly agree).

The usability questionnaire had also 2 open questions: free description of most positive and most negative aspects.

TABLE I
Descriptive Statistics

	Item	Min	Max	Mean	SE	SD
1	Reading the information is easy	3	5	4,34	,106	,602
2	Using the interaction tool is easy	1	5	3,06	,179	1,014
3	Working on the work place is comfortable	3	5	4,28	,121	,683
4	Understanding the system is easy	3	5	4,31	,130	,738
5	Understanding the lesson is easy	3	5	4,38	,117	,660
6	Learning how to operate is easy	3	5	4,31	,138	,780
7	Remembering to operate is easy	3	5	4,41	,126	,712
8	The design of the system is attractive	2	5	3,31	,152	,859
9	The system is fast enough	1	4	2,53	,180	1,016
10	Correcting the mistakes is easy	1	5	3,34	,183	1,035
11	Overall, I found the system easy to use	1	5	3,84	,163	,920
12	Overall, I find the system exciting	3	5	4,16	,111	,628

Our sample was small (n=32), and problems exist both with respect to generalization to the full population as well as with respect to the choice of statistical techniques to be used for the analysis. Therefore descriptive statistics and statistical techniques for compare groups played a relatively important role.

The measures of central tendency and variation are shown in Table I. One item is scored below "neutral" (9) and four are below "agree" (2, 8, 10, and 11). The rest of seven items are scored over 4.00 (agree to strong agree). Individual answers to following items are predominantly in the range "agree"-"strong agree": reading information on the screen, comfort on the workplace, system understanding, lesson understanding, learning how to operate, remembering how to operate and exciting system. There are no disagreements (below 3.00) and there are few neutral answers.

There are some differences in the standard deviations of rating for particular statements, although overall, the patterns are quite similar. For example, the students appeared more certain about items 1, 3, and 5, than about items such as 2, 9 and 10.

B. Gender Analysis

We analyzed if there is a significant difference between the answers given by boys and girls. An independent-sample t-test was conducted to compare the scores for boys and girls. There was no significant difference in scores for boys and girls, with exception the items 8 (The design of system is attractive) and 12 (Overall I find the system exciting). For these two items, Sig. (2-tailed) value is 0.037 and, 0.047 respectively. As these values are less than the required cut-off of 0.05, we conclude that there is a statistically significant difference in the mean scores for boys and girls.

The t test results for the two items (using SPSS 14.0) are presented in Table II.

TABLE II
Gender Analysis

	Gender	N	Mean	SD.	t	df	Sig. (2-tailed)
The design of the system is attractive	Girls	16	3,00	,730	-2,179	30	,037
	Boys	16	3,94	1,063			
Overall, I find the system exciting	Girls	16	3,94	,574	-2,073	30	,047
	Boys	16	4,38	,619			

For other all attributes the Sig. (2-tailed) value is above the required cut-off of 0.05. We conclude that there is *not* a statistically significant difference in the mean scores for boys and girls.

C. Analysis of User Remarks

Usability problems have been identified by analysing the most negative aspects mentioned by students. A summary of usability problems is presented in Table III.

Most frequent was the difficulty to reach each organ with the interaction tool. ("some areas for my position were very unreachable", "I'm not always able to reach everything" or "not every organ is to be reached"). This category of usability problems are related to the selection technique.

TABLE III
SUMMARY OF USABILITY PROBLEMS

Category	Frequency
Selection	31
Interaction tool (paddle)	14
Feedback	14
Discomfort	12
Clarity of sound and writing	7

Second category of negative aspects was the difficulty to use the interaction tool (paddle) which sometimes blocked ("sometimes the cursor isn't moving" or "the program doesn't reacts to my actions sometimes"). Other negative aspects are related to the discomfort provoked by the stereo glasses and the position of the screens.

Overall, usability problems mentioned by students correspond to items in Table I that are scored bellow 4.00.

Most mentioned positive aspects are summarized in Table IV in a decreasing order of their frequency.

TABLE IV
SUMMARY OF POSITIVE ASPECTS

Category	Frequency
Educational support	29
Funny, alike games	19
Attractiveness and comfort	11
Novel, original and interesting	10
Interaction 3D and animation	10
Easy to understand and use	8
Vocal interface and clear explanation	8

Educational support includes aspects like: easy to understand the lesson, stimulating to learn, easy to learn the lesson, usefulness of the demo program, flexibility (potential to do other things) and possibility to repeat the exercise.

The fact that students liked the similarity with a computer game (learning by doing) shows the intrinsic motivation created by the AR technology and the added value of exploiting the need to play, predominant to the age of students ("the system makes me to want to work with it" or "big stimulation of trying to understand the topic"). The 3D interaction and animation are other positive aspects of the AR technology ("the 3D animation raise the interest" or "it is well animated what happens with the food").

Again, the findings are consistent with questionnaire data and show that the AR technology is exciting and increase student's motivation to learn.

V. CONCLUSION AND FUTURE WORK

In this paper, we presented some preliminary results from a formative usability testing. The main findings could be summarized as follows:

- The system has educational value is attractive, stimulating and exciting for students: exercises are alike computing games. The students enjoyed the interaction with 3D objects and the animation using AR techniques.
- Several usability problems have identified. Tools and interaction techniques should be enriched and improved as speed and accuracy. Feedback should be provided when leaving the selection area.

As a concluding remark, we found that formative usability evaluation using both quantitative and qualitative data is a cost-effective support for the user centred design of AR-based educational technologies.

REFERENCES

[1] S. Amershi, N. Arksey, G. Carenini, C. Conati, A. Mackworth, H. Maclaren, D. Poole. (2005) "Designing CIspace: Pedagogy and Usability in a Learning Environment for AI". *Proceedings of ITCSE 2005*. ACM Press, New York. 178 – 182

[2] Ardito, C., M. F. Costabile, A. De Angeli, R. Lanzilotti "Systematic Evaluation of e-learning Systems: An Experimental Validation". *Proceedings of NordiCHI 2006*, 14-18 October 2006. ACM Press. 195-202.

[3] Ardito, C., Costabile, M.F., Marsico, M. de, Lanzilotti, R., Levialdi, S., Roselli, T., Rossano, V. "An approach to usability evaluation of e-learning applications." *Universal Access in the Information Society, 4* (2006): 270-283. Springer

[4] Augmented Reality in School Environments – FP6 027039 http://www.arise-project.org/

[5] Azuma, R. A Survey of Augmented Reality. PRESENCE: Teleoperators and Virtual Environments, Vol. 6, No. 4, pp. 355-385, 1997.

[6] Barajas, M. and Owen, M., Implementing virtual learning environments: looking for holistic approach. *Educational Technology and Society*, 3, 2000.

[7] Bogen, M., Wind, J. and Giuliano, A.: "ARiSE - Augmented Reality in School Environments." Nejdl, W. & TochTermann, K. (Eds.): *Proceedings of EC-TEL 2006. LNCS 4227,* Springer 2006. pp. 709-714

[8] Hix, D., Gabbard, J., Swan, E., Livingston, M., Herer, T., Julier, S., Baillot, Y. & Brown, D.: A Cost-Effective Usability Evaluation Progression for Novel Interactive Systems, In *Proceedings of the Hawaii International Conference on Systems Sciences*, January 5-8, 2004, Big Island, Hawaii.

[9] Kukulska-Hulme, A. & Lesley Shield (2004), *The keys to usability in e-learning websites,* Networked Learning Conference 2004, Lancaster.

[10] Liarokapis, F., Mourkoussis, N., White, M., Darcy, J., Sifniotis, M., Petridis, P., Basu, A. & Lister, P.: Web3D and augmented reality to support engineering education in *World Transactions on Engineering and Technology Education Vol.3, No.1,* 2004.

[11] Nielsen, J. Usability Engineering. Academic Press. 1993.

[12] Nokelainen, P. „An empirical assesment of pedagogical usability criteia for digital learning material with elementary school students". *Educational Technology & Society, 9 (2)*, 2006. 178-197.

[13] Silius, K. & Tervakari, A-M. "The usefulness of web-based learning environments. The Evaluation Tool into the Portal of Finnish Virtual University." *International Conference on Network Universities and e-learning*. 8-9 May 2003. Valencia. Spain.