Summary of Usability Evaluations of an Educational Augmented Reality Application

Hannes Kaufmann[1], Andreas Dünser[2]

[1] Interactive Media Systems Group, Vienna University of Technology
Favoritenstrasse 9-11/188/2, A-1040 Vienna, Austria
kaufmann@ims.tuwien.ac.at

[2] HIT Lab NZ, University of Canterbury, Private Bag 4800
8140 Christchurch NZ
andreas.duenser@hitlabnz.org

Abstract. We summarize three evaluations of an educational augmented reality application for geometry education, which have been conducted in 2000, 2003 and 2005 respectively. Repeated formative evaluations with more than 100 students guided the redesign of the application and its user interface throughout the years. We present and discuss the results regarding usability and simulator sickness providing guidelines on how to design augmented reality applications utilizing head-mounted displays.

Keywords: augmented reality, usability engineering, formative evaluation, geometry education

1 Introduction

Our work is based on the educational Augmented Reality (AR) application Construct3D [1-3]. This system deploys AR to provide a natural setting for face-to-face collaboration of teachers and students. The main advantage of using AR is that students actually see virtual three dimensional objects. With traditional methods students have to rely on 2D sketching or calculating and constructing objects using pen and paper or CAD software. Direct manipulation and dynamic interaction with virtual 3D objects using tangible interaction devices are key features of Construct3D. In our standard setup users are wearing a see-through head-mounted-display; a pen and a panel are used for direct interaction in 3D space. Head, pen and panel are fully tracked in 3D which allows users to walk around objects and to view them from different perspectives (Fig. 1).

By working directly in 3D space, complex spatial problems and spatial relationships may be comprehended better and faster than with traditional methods. Our system utilizes collaborative AR as a medium for teaching, and uses 3D dynamic geometry to facilitate mathematics and geometry education.

Over the course of 6 years Construct3D has been developed, improved, tested and evaluated with more than 100 students in over 500 teaching lessons. Pedagogical theories such as constructivism and activity theory influenced the design of the collaborative educational AR hardware setup and content design. Technical details and pedagogical uses of Construct3D (including teaching content) have been published by the first author before [2-4].

Fig. 1. Students working with Construct3D.

The development process of Construct3D resembles the usability engineering methods of virtual environments suggested by [5]. The first informal evaluation in 2000 helped to compile a detailed user task analysis whereas expert guideline-based evaluations occurred numerous times during the development process. Visiting teachers and researchers evaluated the system and provided useful feedback. Two formative evaluations in 2003 and 2005 had a big impact on the design and development of Construct3D. In this paper we summarize three usability evaluations conducted in 2000, 2003 and 2005 and will present the lessons learned.

2 Construct3D

Construct3D is based on the Studierstube AR system [6]. It promotes and supports exploratory behavior through dynamic 3D geometry. A fundamental property of dynamic geometry software is that dynamic behavior of a construction can be explored in real time by interactively moving individual defining elements such as corner points of a rigid body. Users can see which parts of a construction change and which remain the same. The histories of constructions as well as dependencies between geometric objects are maintained. Experiencing what happens under movement facilitates better comprehension of a particular construction and geometry in general.

The menu system is mapped to a hand-held tracked panel called the personal interaction panel (PIP) [7]. The PIP (Fig. 2) allows the straightforward integration of conventional 2D interface elements like buttons, sliders, dials etc. as well as novel 3D interaction widgets. Passive haptic feedback from the physical props guides the user when interacting with the PIP, while the overlaid graphics allows the props to be used as multi-functional tools. Students can for instance position written notes onto the tablet which might help them during their work in the virtual environment.

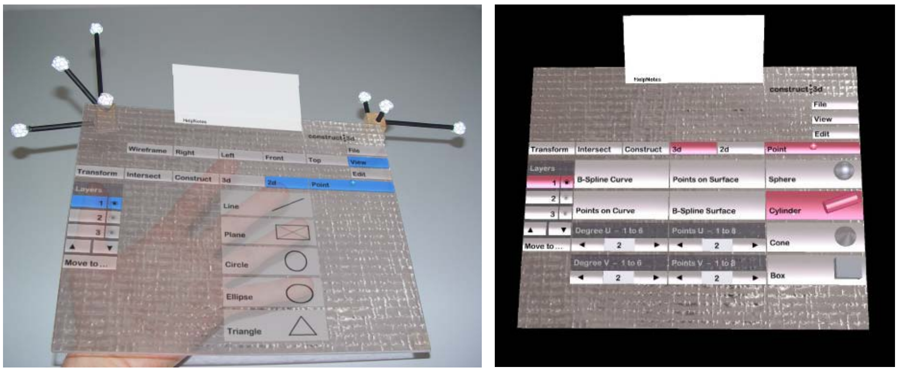

Fig. 2. Left: Menu system of Construct3D displayed on the PIP. In a help-box (on top) further details and help on application features are provided. Right: 3D submenu displayed for the user working with the red color scheme.

All construction steps are carried out via direct manipulation in 3D using a stylus tracked with six degrees of freedom. In order to generate a new point the user clicks with his pen exactly at the location in 3D space where the point should appear. Users can easily switch between point mode (for setting new points) and selection mode (for selecting 3D objects).

Desktop CAD systems typically have a very steep learning curve and offer an abundance of features in deeply nested menus. For Construct3D we focused on a simpler menu system, which is easy to learn and intuitive to use. In addition we accommodated to the fact that menu widgets seen through a HMD need a certain size in order to be usable. Organizing the functions proved difficult under these conditions as the number of program functions increased over time. We finally organized the menu – according to a user task analysis, experts' guidelines and experience by logic grouping of functionality – into five submenus accessible via tabs (Fig. 2), with frequently used functions being visible all the time. This provides relatively quick access to all program functions. The menu concept is similar to that used in traditional desktop CAD menu systems known by many students, while avoiding excessive interface modes.

Hardware Setups. The standard immersive setup used for Construct3D supports two collaborating users wearing stereoscopic see-through head mounted displays (HMDs) (see Fig. 1) providing a shared virtual space. The users interact with the system using pen and pad props (Fig. 2). Both users see the same virtual objects as well as each others' pens and menu systems which provides a global shared space. In addition it allows users to help each other (i.e. with the menu system) if necessary. Position and orientation of head and hands are tracked using a 4-camera infrared-optical tracking system. In a co-located setup - such as the one used for our evaluations - one dedicated host with two graphic ports renders stereoscopic views for both users.

3 Usability studies

We report and compare a first informal user study and formative usability studies completed in 2003 and 2005. Based on feedback from many trials with high school students and a first informal evaluation in 2000 [8] we continuously improved Construct3D over a course of 5 years.

All usability enhancements were conducted with the intention of improving collaborative learning and teaching. As usability can only be improved in accordance with users' needs and application specific strengths and weaknesses, the guidelines mentioned here cannot be applied directly to other applications without careful adaptation.

3.1 1st Informal Evaluation – 2000

In our first evaluation [8] with 14 students we observed the students' interaction with the system. We obtained very positive and encouraging feedback and a number of problems were pointed out. During the evaluation it was gratifying for us to see users work with Construct3D in a very constructive manner. They did not need a long introduction to the system but applied their experience with 2D user interfaces to the 3D interface. After completing the task, some walked around the objects, viewing them from different sides or got down on their knees and looked at the scene from below. Half of the students felt that working with Construct3D for the first time was easier than their first experience with a desktop CAD package.

Hand-eye coordination showed to be very difficult when spotting a point accurately in 3D space without haptic feedback or constraints. All students reported problems with setting points at given coordinates. As a consequence we implemented raster and grid functions. About constructing in VR, students especially liked walking around and inside objects, the "playful" way of constructing, and that spatial relationships and complex three dimensional designs are directly visible. The clear structure of Construct3D's menu system and the audio help system were mentioned positively.

At that time Construct3D was still a static modeling tool and did not provide dynamic features. Insights gained from the first evaluation (i.e. the difficulty for highly accurate 3D interaction) and the understanding that students would educationally benefit from 3D dynamic geometry encouraged us to change Construct3D into a dynamic 3D geometry application.

3.2 2nd Evaluation Study - 2003

In 2003 we conducted a study based on interviews and the standardized ISONORM 9241/10 usability questionnaire [9]. We designed a number of training exercises that fit the Austrian descriptive geometry curriculum of 11th and 12th grade [4]. Using Construct3D, 15 high school students (9 male, 6 female) worked on these exercises with the aid of their teachers. All students attended geometry classes (descriptive geometry) since the beginning of grade 11. Each of them participated in 5 training sessions lasting 6 hours. Our main objective was to assess the usability of our system and its potential as an educational tool for real high school work. At the end of all training sessions students had to answer an ISONORM usability questionnaire. Two questions regarding self-descriptiveness of the application had to be removed since they were related to desktop applications only. Afterwards students answered general questions regarding user acceptance, user behavior, technical requirements and organizational aspects.

Results. A closer look at the data (Figure 2) reveals that the categories "suitability for learning" and "suitability for task" received the highest rating which is very important in this context. In our opinion the highest priorities for an educational application that complies with pedagogic theories such as constructivism are that it (1) is easy to use and requires little time to learn, (2) encourages learners to try new functions and (3) can be used consistently and is designed in a way that things you learned once are memorized well. These are exactly the items that students rated very high. Almost all students reported that they could imagine using the current version of Construct3D in high school or university education.

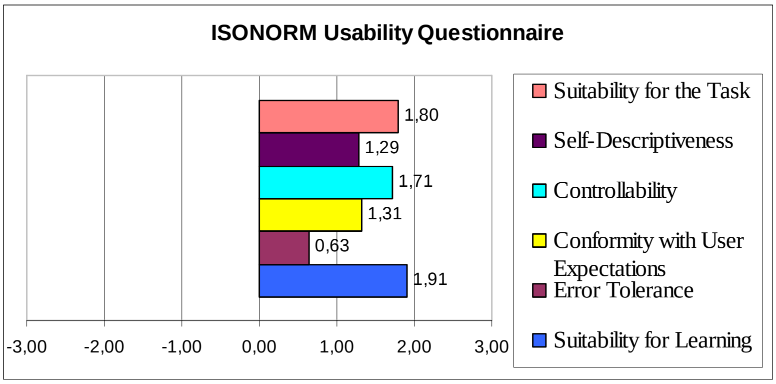

Fig. 3. Results of the ISONORM [9] usability questionnaire in 6 categories

The categories "self-descriptiveness" and "conformity with user expectations" got lower ratings than the rest. Self-descriptiveness of Construct3D was improved by adding better labeling and a help-box on the panel in order to explain all menu items.

As a result of this usability study the user interface was completely redesigned. The menu system was restructured (Fig. 2) to make features that are used most frequently easily accessible. In addition the visual design of geometric objects was enhanced considerably. The purpose of visual design of objects constructed by the user is to support the user's understanding of a construction. Unlike desktop visualization of the same content, using stereoscopic see-through HMDs requires to deal with limited contrast, resolution and viewing angle. Moreover, the system should present scenes of high depth complexity in a clear way, providing an improved insight into the construction. Among the techniques employed in Construct3D to support these goals are the use of transparencies for geometric objects to allow students to see inside objects (Fig. 1), consistent color coding to allow distinguishing between multiple users' contributions (which is especially important in distributed remote teaching scenarios), separation into layers to support semantic structuring of a construction, and automatic previewing of new objects. Details of the improvements are given in [3].

3.3 3rd Evaluation Study - 2005

In the 2005 evaluation 47 students were solving tasks with Construct3D in AR while another group of 44 students solved the same geometric problems with an educational desktop application called CAD3D [10] (which is used in Austrian high schools). Participants were Austrian high school students aged between 16 to 19 years (M = 17.49, SD = .79; 44 (48.4%) male and 47 (51.6%) female). Students attended 6 training sessions which lasted 45 minutes with one week pause in between. In both groups a tutor supervised two students working on the geometry tasks. The tutors explained the tasks to the students and supported them if they needed help.

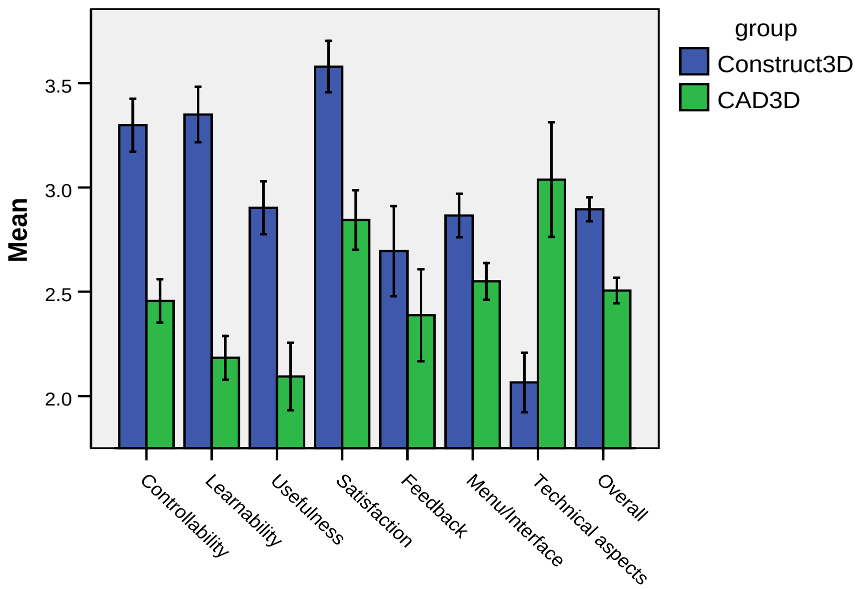

Fig. 4. Usability ratings of students working with Construct3D and CAD3D (4-point Likert scale; 1-min, 4-max = best; error bars ± 1.96 * standard error)

To assess usability we adapted questions of 8 established usability questionnaires to develop a questionnaire (7 scales (see Fig. 4); 28 questions in total) better suited for the range of applications tested. The questions were taken from the Questionnaire for User Interface Satisfaction, Perceived Usefulness and Ease of Use, Purdue Usability Testing Questionnaire, Computer System Usability Questionnaire, Practical Heuristics for Usability Evaluation (all at [11]), Software Usability Measurement Inventory [12], SUS [13] and the ISONORM [9] usability questionnaire.

Results. The analysis of the usability questionnaire showed that students using Construct3D gave higher ratings ($p < .01$) for all categories (Controllability, Learnability, Usefulness, Satisfaction, Feedback, and Menu/Interface) except technical aspects (e.g. robustness) than students using CAD3D. This indicates that the AR based geometry education application Construct3D is a highly usable system which - from a usability perspective - has several advantages over the traditional desktop based application. Especially user satisfaction, learnability and controllability got high ratings. However the low ratings for technical aspects suggest that there are still some issues regarding technical robustness that have to be addressed. Infrequent system crashes and minor technical problems can reduce motivation of participants and usability of the system.

Comparing the results of the 2003 and 2005 evaluations illustrates that conformity with user expectations (2003) / satisfaction (2005) was improved throughout the years. Suitability for the task got quite high ratings in the 2003 evaluation. In 2005 students rated usefulness, the equivalent scale, somewhat lower. In the 2005 evaluation a more extensive training setup was realized and thus students worked on a broader variety of geometric problems (e.g. problems used in standard school curriculum). Hence, this result may indicate for which kind of geometric problems Construct3D is a suitable educational tool. In both formative evaluations its strengths became obvious. Construct3D should mainly be used for teaching content which utilizes 3D dynamic geometry or requires the visualization of abstract problems. In addition these are areas that are hardly covered by other educational applications.

We also asked the students other questions in order to get more detailed feedback on the training task and setup. Analyzing the students' answers to these questions may help to refine our system setup and further adapt it to users needs. Table 1 shows the preferred training setup of students using Construct3D and CAD3D.

Table 1. How would you prefer to work with Construct3D / CAD3D

	Construct3D	**CAD3D**
2 students, one tutor (like in the training sessions)	80.95%	86.00%
1 student, one tutor	9.52%	4.65%
2 students, without tutor	4.76%	2.33%
alone	4.76%	4.65%

There were no significant differences regarding the preferred training setup between students working with Construct3D and CAD3D. Most of the students liked the setup we used for the trainings in our study: 2 students working with one tutor.

Regarding the potential use of Construct3D in educational institutions we asked the students if they would like to use Construct3D in school in a setting similar they worked with (1 to 2 students) given the technical equipment would be affordable for schools. The majority of students would like to use Construct3D in school (yes = 64.44%, rather yes = 26.67%); 8.89% would rather not like to use the system in schools. Students' comments on the potential problems of using Construct3D in schools were mainly concerned with lack of finances and the robustness of hardware and software.

4 Simulator Sickness

As described earlier, Construct3D requires users to wear a HMD. In the second evaluation study (2003) some of the students reported negative side effects after working in the virtual environment, a condition known as simulator sickness, which is similar to motion sickness [14]. One female student reported headache and eye strain after 20 minutes of work in the virtual environment but did not stop working and wanted to use Construct3D again. In retrospect we know that our evaluation sessions lasting one hour were simply too long for continuous work with a HMD. Since negative side effects are a general potential problem when working with HMDs and influence the user's subjective experience of a VR/AR environment considerably they are relevant to all VR/AR applications that use these displays. We identified some possible reasons of such negative side effects that may be relevant to our virtual environment such as accommodation problems, low frame rate, lag or bad fitting helmets. If not taken into account, symptoms experienced by users affected by simulator sickness can drastically diminish usability of a system [15].

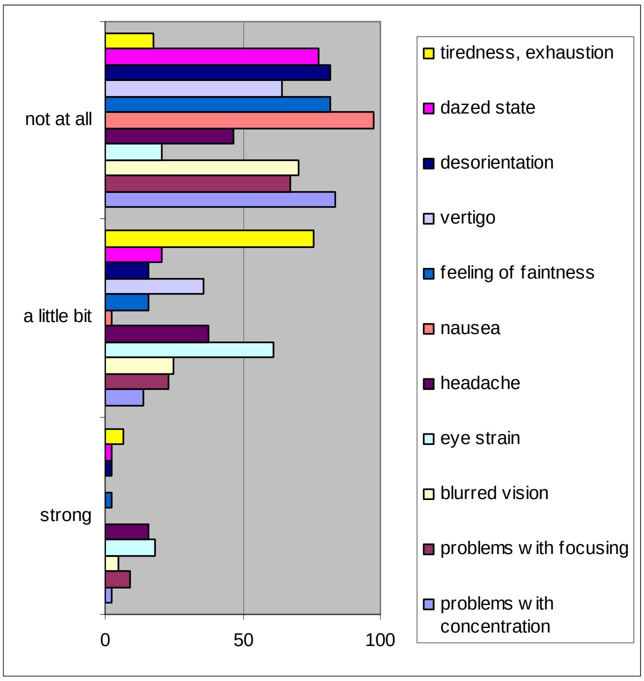

Fig. 5. Percentage of users reporting a specific symptom is shown (0% = reported by no user; 100% = reported by all users).

In order to minimize the chance of users suffering from symptoms of simulator sickness we limited training sessions to a maximum of 45 minutes in our third evaluation study (2005). Furthermore we replaced the hard plastic helmet (Fig. 1, left) which caused pressure on some students' forehead or even headache with a relatively lightweight bicycle crash helmet (Fig. 1, right). Students also were asked to take a rest when they felt the need to. After they finished the training sessions with Construct3D we asked them to which extent they actually did experience specific symptoms related to simulator sickness (questionnaire; 11 questions). Fig. 5 shows the percentage of participants having experienced a specific symptom 'not at all', 'a little bit' or 'strong' during or while having worked with Construct3D.

75.56% of the 47 participants felt a moderate amount of tiredness or exhaustion and 61.36% reported a little bit of eye strain. There were also some participants who reported having experienced some headache (37.78%) and vertigo (35.56%). Most of these symptoms may be related to the use of a HMD. Thus although we limited training time there still seem to persist issues with respect to some simulator sickness symptoms, especially exhaustion and eye strain. However in general most of the participants did not report having experienced severe problems.

In accordance with our observations and other studies we recommend limiting HMD usage to 20-30 minutes per session. Based on our experience image quality of HMDs but especially lag and quality of tracking data contribute most to the reported effects.

5 Conclusion and Future Work

In this summary of usability evaluations we describe how we managed to improve usability of Construct3D iteratively. We gradually adapted, reconfigured and redesigned hard- and software and integrated new interaction techniques and interfaces according to our observations and user feedback. A number of studies report that cognitive overhead in mastering the interface can hinder training and learning of the task [16]. Especially in educational applications it is of utmost importance to focus students' attention on the actual task and to reduce cognitive overhead needed to use the application. This motivated us to put a lot of time and effort into interaction and user interface design. We gained valuable results from the evaluations which helped us to create a more usable AR-based learning environment with improved user satisfaction.

In our latest evaluation we found that the usability of Construct3D was rated higher than the usability of a desktop based geometry education application. This may be due to the more intuitive workflow when working on 3D tasks. However there are still technical issues (e.g. software robustness) that have to be solved in order to improve usability even further. Especially problems related to the use of HMDs and tracking latency need careful thought. Thus at this stage we recommend to limit usage times of head mounted displays in immersive training setups. For an educational application such as Construct3D we envision its integration into courses; therefore temporally limited usage is very reasonable in this context.

Developers of AR-based applications face specific hard- and software related issues that are different from those of desktop based GUI or WIMP design. No set of common design guidelines exist yet that would facilitate or streamline the development of easy to use AR systems [15].

Regarding future work we plan to use Construct3D as a tool for evaluating various aspects of virtual learning environments in our future research including a comprehensive pedagogic evaluation, studying e.g. teaching styles/methodology or transfer of learning to tasks in the real world.

References

1. Kaufmann, H., Schmalstieg, D.: Mathematics and geometry education with collaborative augmented reality. Computers & Graphics **27** (2003) 339-345
2. Kaufmann, H.: Geometry Education with Augmented Reality. Ph.D. Thesis. Vienna University of Technology, http://www.ims.tuwien.ac.at/research/construct3d/videos.php (2004)
3. Kaufmann, H., Schmalstieg, D.: Designing Immersive Virtual Reality for Geometry Education. Proceedings of IEEE Virtual Reality Conference 2006, Alexandria, Virginia, USA (2006) 51-58
4. Kaufmann, H., Papp, M.: Learning Objects for Education with Augmented Reality. Proceedings of EDEN 2006 (European Distance and E-Learning Network) Conference, Vienna (2006) 160-165
5. Hix, D., Gabbard, J.L.: Usability Engineering of Virtual Environments. In: Stanney, K.M. (ed.): Handbook of Virtual Environments - Design, Implementation, and Applications. Lawrence Erlbaum Associates, Mahwah, New Jersey (2002) 681-699
6. Schmalstieg, D., Fuhrmann, A., Hesina, G., Szalavári, Z.S., Encarnacao, L.M., Gervautz, M., Purgathofer, W.: The Studierstube augmented reality project. Presence - Teleoperators and Virtual Environments **11** (2002) 33-54
7. Szalavári, Z.S., Gervautz, M.: The Personal Interaction Panel - A Two-Handed Interface for Augmented Reality. Computer Graphics Forum **16** (1997) 335-346
8. Kaufmann, H., Schmalstieg, D., Wagner, M.: Construct3D: a virtual reality application for mathematics and geometry education. Education and Information Technologies **5** (2000) 263-276
9. Prümper, J.: Der Benutzungsfragebogen ISONORM 9241/10: Ergebnisse Zur Reliabilität und Validität. In: Liskowsky, R. (ed.): Software-Ergonomie '97, Stuttgart (1997)
10. Stachel, H., Wallner, J., Pfeifer, M.: CAD-3D für Windows. http://www.geometrie.tuwien.ac.at/software/cad3d/ (2003)
11. Perlman, G.: Web-Based User Interface Evaluation with Questionnaires. http://www.acm.org/~perlman/question.html (accessed Feb. 15th 2006)
12. Kirakowski, J., Corbett, M.: SUMI: the Software Usability Measurement Inventory. British Journal of Educational Technology **24** (1993) 210-212
13. Brooke, J.: SUS: a "quick and dirty" usability scale. In: Jordan, P.W., Thomas, B., Weerdmeester, B.A., McClelland, I. (eds.): Usability Evaluation in Industry. Taylor and Francis, London (1996) 189-194
14. LaViola Jr., J.J.: A discussion of cybersickness in virtual environments. ACM SIGCHI Bulletin **32** (2000) 47-56
15. Dünser, A., Grasset, R., Seichter, H., Billinghurst, M.: Applying HCI principles to AR systems design. MRUI'07: 2nd International Workshop at the IEEE Virtual Reality 2007 Conference, Charlotte, North Carolina, USA (2007)
16. Dede, C., Salzman, M.C., Loftin, R.B.: ScienceSpace: Virtual Realities for Learning Complex and Abstract Scientific Concepts. Proceedings of IEEE VRAIS '96 (1996) 246-252