

USABILITY OF THERAPIST'S USER INTERFACE IN VIRTUAL REALITY EXPOSURE THERAPY FOR FEAR OF FLYING

Lucy T. Gunawan &
Charles van der Mast
Delft University of Technology
Mekelweg 4 2628 CD Delft
the Netherlands
Email:
c.a.p.g.vandermast@ewi.tudelft.nl

Mark A. Neerincx
Delft University of
Technology, Mekelweg 4,
2628 CD, Delft, & TNO
Human Factors, Soesterberg,
the Netherlands
Email: neerincx@tm.tno.nl

Paul Emmelkamp &
Merel Krijn
University of Amsterdam,
Roetersstraat 15,
1018 WB Amsterdam
the Netherlands
Email: p.m.g.emmelkamp@uva.nl

KEYWORDS

Virtual reality, human-computer interaction, usability, phobia treatment.

ABSTRACT

Delft University of Technology has implemented an experimental virtual reality system for treating fear of flying. The patient sits in an airplane chair equipped with vibration devices. By wearing a head-mounted device, the patient is immersed in the virtual world (patient's user interface), sitting in the computer-generated cabin of a virtual airplane and experiencing various aspects of flying. The therapist takes control of what will be experienced by the patient, by using the therapist's user interface. In order to perform comparative studies with the system the patient's user interface was made optimal first for effective treatments. Now the therapist's user interface should be given higher usability. After analyzing the current user interface and its deficiencies, some improvements were carried out. These improvements include a new design of therapist's user interface, including database support and reporting tools, and adding some important features in the patients user interface: lightning simulation, time of the day background, cabin density control, flying destinations, attendant announcements, airplane rolling,. The improvements were evaluated with a usability analysis. Interviews and laboratory experiments were done to measure the usability. And as we expected, the results gave positive indicators leading to improvements of the system usability – as seen by the therapist.

INTRODUCTION

Recent research (e.g. Emmelkamp et al. 2002) has proved that virtual reality technology (VR) can be implemented in clinical therapy.

One of the virtual reality projects at Delft University of Technology is virtual reality exposure therapy (VRET) for treating fear of flying (Schuemie 2003). This project is a collaboration between two disciplines, Psychology, brought in by University of Amsterdam and VALK foundation in Leiden, and Human-Computer Interaction (HCI), brought in by Delft University of Technology. Delft University of Technology is in charge in the technical part for this project, developing the VRET system. The system has two main user interfaces, one for the therapist to control the

session and one for the patient to experience flying. This experimental system was deployed successfully during experimental therapy sessions in comparative and other studies for acrophobia, claustrophobia and agoraphobia (Schuemie 2003).

Our goal is not only to implement this system but also to support clinical therapy professionally. Therefore, the intention of this research is to improve the usability of the system, both for the therapist's and the patient's user interface. Schuemie (2003) worked mainly on the usability of the patient's interface by parameters of presence and locomotion. In this paper, we only discuss the therapist's user interface improvements. The complete reference to patient's user interface is described by Schuemie (2003).

Our first step was to collect feedback by the few therapists who used the system during the comparative studies on treatment in vivo and in VR. By using the information gathered, we developed some improvements. The therapists participated in every step of system development.

THE SYSTEM

The patient sits in a real airplane seat equipped with a bass loudspeaker in the seat to simulate vibration in the airplane. By wearing the head-mounted device (HMD), the patient is immersed in a virtual world, enters the computer-generated cabin of the virtual airplane and will experience various aspects of flying controlled by the therapist who is nearby in the same room. The patient is exposed to flying situations such as: sitting in a standing still airplane, taxiing on the runway, taking off, flying in good weather, flying in bad weather and landing. During therapy, therapist can see and hear patient's experience during the virtual flight. The therapist works most of the time using the therapist's computer where he/she can control the VR world, monitor the patient regularly and check the level of fear experienced by the patient, see Appendix A.

The first version of this system was built in 1999 by Schuemie, see (Schuemie & Van der Mast 2001). In this paper are described the details of hardware and software specifications.

REQUIREMENTS

From the interviews with the therapists was found that they need to have more control over the aircraft and what is happening during the flight. In the old therapist user

interface (see Appendix B) an overview of all the functionality is shown. Input was given via keyboard, mouse and joystick. The new requirements should be important to be able to give a more flexible treatment reacting to individual characteristics of the patients. The therapists asked e.g. for more destinations (in the old user interface only Milan), flying during day or night, more bad weather control, control of flap wings and rolling.

METHOD

The adaptation of the usability evaluation model proposed by (Scriven 1967) quoted by (Rosson & Carroll 2002), is used as our research methodology (Figure 1). The model distinguishes between formative and summative evaluations. The goals of formative evaluation are to identify the design aspects that can be improved, to set priorities, and to provide guidance in how to make changes to a design. This evaluation is conducted during the design and development process. The summative evaluation goals are to measure quality; to evaluate a design result whether the system has met its usability objectives and it is conducted at the end of development process. This model can be seen as an iterative process that the current system can be evaluated as if it still is in the design process, although it was finished before.

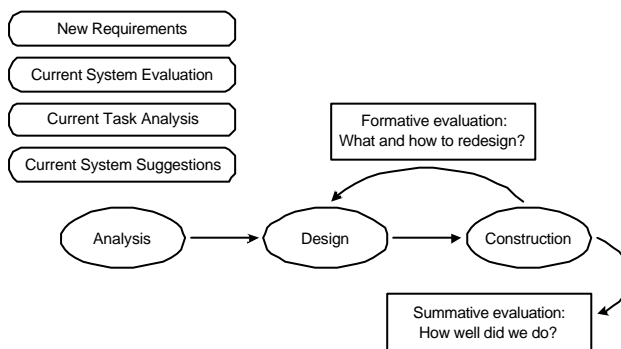


Figure 1: Research methodology

The analysis process had some inputs such as task analysis of the current system, new requirements gathered by interviewing the therapists, evaluation of the current system and some suggestions from the research partners at University of Amsterdam. In the design phase, we designed two kinds of user interface (UI), the virtual world for the patient and the user interface control for the therapist. Our main concern in this study was to improve the usability of the therapist's user interface, but sometimes we have to enhance the patient's UI in order to extend the therapist's UI. New features were added, improvements were carried out, and a new therapist user interface was implemented. The evaluation of the therapist's UI is measured in terms of usability, i.e. effectivity, efficiency and satisfaction (Rosson and Carroll 2002). We used a mediated evaluation which is a mix between analytical and empirical method. The analytic evaluation is done early and during the design process. The result of this analysis is used to motivate and develop materials for empirical evaluations. Heuristic evaluation as an inspection method was done and also the

ten general heuristic guidelines by Nielsen (1993) were taken into account. The empirical method is done by a user evaluation experiment. Because we have only a very limited number of professional therapists for treating fear of flying we added other persons to do the same treatment/job. From Neerinx et al. (2001) we know that this may deliver valid results. The therapists tested the system directly with a user/patient in the virtual airplane. Some specific tasks were given to the therapist to complete. Information was gathered, such as the observation protocol, performance time, errors and subjective evaluation. The subjective evaluation was acquainted by using usability questionnaires and interviews.

DESIGN AND DEVELOPMENT OF THE NEW USER INTERFACE

Based on the task analysis by Schuemie (2003), the task model of current in vivo therapy for phobia treatment was formed. The main goal of each therapy is to cure the patient. During the exposure, the therapist determines patient's fear by exposing manipulating stimuli to patients, and changing it when needed to adjust patient's fear. The therapist responds to each question the patients might have. This solves any ambiguity patients might have. Responding or answering the patient's question might not have contribution to cure the patient directly, but at least it will facilitate patients in performing their tasks. Patients believe that by following the therapist's instruction, they can get rid of their fear. People with phobias have strong tendency to avoid fearful situations. This conflicts with therapist's instruction. To resolve ambiguity in therapy, the patient sometimes need to inquire about certain matters.

As the other input to the analysis, there were also some suggestions in the several areas where usability can be improved such as providing the therapist with cognitive artifacts representing the historical patient's score over the therapy and increase the learnability and memorability of using the system.

Though all new requirements urged to be added, some consideration was taken, and we could not implement them all. A new UI for the therapist was designed, some features to the world were added such as: lightning, possibility of flying during different time of the day (morning, day, afternoon, and night), possibility to change the cabin's passenger density, possibility to fly to another destination, possibility to choose the voice of pilot and purser, possibility to roll the airplane during the flight, possibility to dim the cabin's light and the most important one is the feature of database, the possibility to save and print historical data of the patient with its Subjective Unit Discomfort (SUD)s artifacts.

The overview of the old therapist's user interface and the new one can be seen in appendix B and C.

RESULTS

The evaluation phase took place in the end of system development. The evaluation goal was to evaluate the usability of the therapist UI whether the new features added and changed showed significant improvement. Thus we formulated our hypothesis as follow: *The "improvements" in Therapist UI are increasing the usability of the system.*

We did one experiment with sixteen pairs of patients and therapists (32 participants), five of them were real therapists, and the rest were students. There were two therapy sessions for each therapist and patient pair, one using the old system (System A) and one using the new improved system (System B). The order was at random. Eleven students were trained for this experiment as therapists (most of them never used our system before) and five real therapists were asked to do therapy sessions. To give more objective judgment of the two systems, none of them was informed which the old system was and which the new improved system was. Each session took about twenty minutes and there were a small break between the sessions. Detailed therapy session tasks was given to the therapist. It included instruction to load the correspondence world, fill patient and session information, gradually expose the patient to the flying sequences in virtual world and end the simulation. Time elapsed was recorded during each task and what the therapist done was monitored and noted such as mistakes done by therapist, questions, and assistance needed. Each therapist had to fill in the usability questionnaire after each session. After finishing the therapy sessions, the therapist were asked about general remarks, comments, suggestion and general comparison about two systems. Another extended subjective evaluation with real therapist also was done for gathering information that is more authentic.

Usability Questionnaire

The reliability analysis for the usability questionnaire was performed. Cronbach alpha was 0.9254 (N of cases=32, N of items=27), showed that responses have a really good internal consistency.

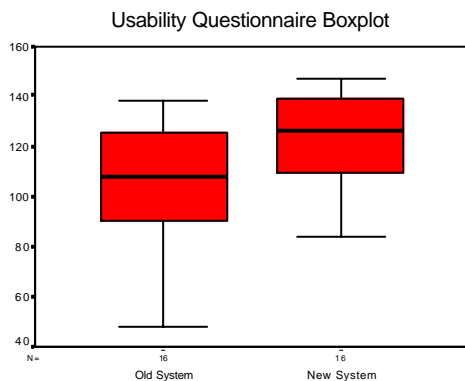


Figure 2: The box plot of the usability questionnaire for the old and the new system

We analyzed our data using ANOVA for repeated measures, because in our experiment the same patient took part in two sessions (the old and new system). The ANOVA for repeated measures show a significant difference in the total score between two systems: $F=12.321$, $p=0.004$. This significant difference should be verified; which one had the higher or lower usability? By calculating the means of questionnaire for old and new system, we drew a box plot as seen in Figure 2. This figure shows us that the means of the new system (122.75) is higher than those of the old system (105.44). Thus, by this result we accepted our

hypothesis that the improvements made in therapist UI increase the usability of the system. There were no significant differences between the groups who tried the systems in a different order: the old system for this first session followed by the new system or the new system for this first session followed by the old system. ($F=1.310$, $p=0.275$). There were no significant different results between the real therapist and student as therapist when they fill in the questionnaire. ($F=0.207$, $p=0.658$).

Table 1: Average scores (and standard deviation) of the additional usability questions (n=16) regarding new features of the therapist UI, scale from 1 to 7.

Element	Usefulness	Ease of use
Flight Plan Control		4.9375(1.3401)
Cabin Control		4.8750(1.0247)
Roll Control		4.8125(1.1087)
Flight View		5.3125(1.0145)
Print Function	5.1250(1.2583)	4.8750(1.5000)
Timer Feature	4.5000(1.3166)	
Simulation Control		4.9375(1.5262)

Additional usability questions filled only for the new system to evaluate the new features had a reliability alpha of 0.8216 (N of cases=16, N of items=8). It means that these eight additional questions had a good internal consistency. The results of the additional questionnaire are displayed in Table 1, which shows that all the new features were evaluated positively. They were very useful and/or easy to use. Thus, by these results we add our hypothesis to include proof that the new added features are useful and easy to use.

We found also a significant correlations between usability of the old and the new system (Pearson Correlations=0.727, $p=0.01$). A higher score in the usability questionnaire of the old system tend to paired with a higher score in the usability questionnaire in the new system.

Performance Time and Error

Performance time and error were measured during the experiment, in every task given. The first task (task 1 in Figures 3) was to load the virtual environment. The second task (task 2 in Figures 3) was to fill the session information, such as: patient, therapist and session number. The third task (task 3 in Figures 3) was to gradually expose the patient in the virtual world.

Table 2: Means of tasks completion time (in seconds)

Participant	Old System	New System
Real Therapist	1090.64	1344.50
Student as Therapist	969.06	1209.86

We could not really compare the time completion between the old and new system, because they are different worlds with different added features. The new system always takes longer time to complete everything, because of new added features. What we can see here is the comparison of time completion between the real therapists and students as therapist. The student as therapist tends to complete the task

faster than the real therapist, as you can see in *Table 2*. This is maybe because students don't really know the pace of therapy session.

Errors in our system were defined as errors made by the therapist during the therapy sessions, and when assistance was needed. The comparison of error rate for both systems can be seen in *Figure 3a, 3b, 3c*. From the graphics we can see that the error rate for the new system is better than that of the old system. We can also compare the means of error rate between real therapist and student as therapist. We can see here the difference, that students make more errors than the therapists.

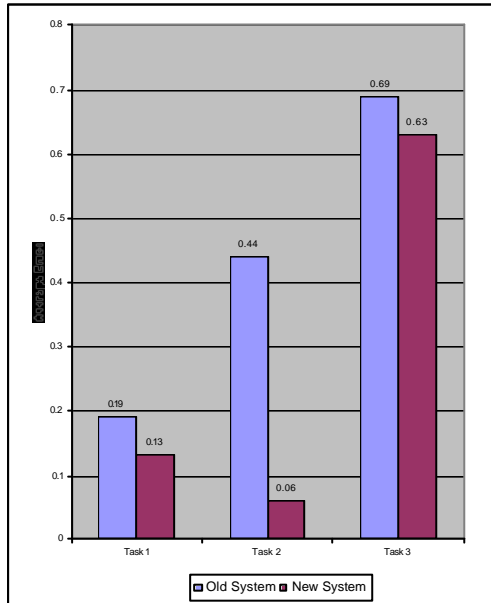


Figure 3a: Bar chart comparison of average error rate for old and new system

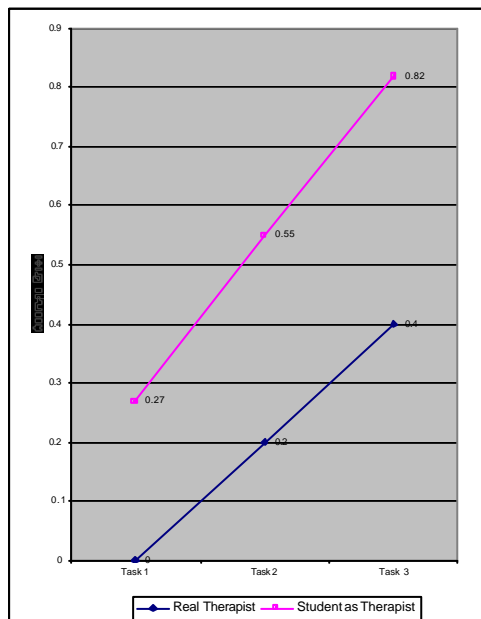


Figure 3b: Old system comparison of average error by real therapist and students as therapist

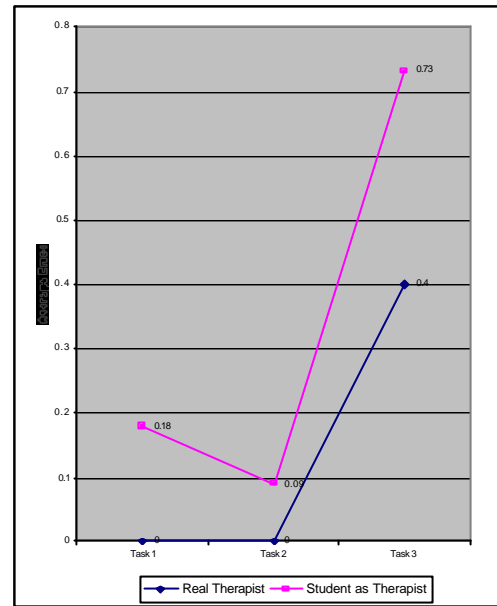


Figure 3c: New system comparison of average error by real therapist and students as therapist

Therapist Subjective Evaluation

Five therapists were given more questionnaires after conducting two sessions. The results of the first five questions are summarized as in *Table 3*. The roll control was not used often, and it was not too easy to use either. The roll control is used during flying stage. It rolls the airplane so the horizon will slightly leaning in the patient's view. It was unclear in the therapist's UI when the roll control can be used. The roll control should represent a continuous control process rather with a discrete button.

Table 3: Average scores (Standard Deviation) of the therapist subjective evaluation, scale from 1 to 5.

Element	Frequency of use	Ease of use	Usefulness
Roll Control	2.20(1.3038)	3.60(1.9494)	-
Bad Weather Control	3.60(1.6733)	4.80(0.4472)	-
Flight View	4.00(1.7321)	-	4.80(0.4472)
Timer feature	-	-	4.80(0.4472)
Print Function	-	-	4.40(0.8944)

The therapists did not too often use the bad weather control but the easiness of this control was evaluated very positively. The new feature of flight view was evaluated positively both for the frequent use and usefulness, this allowed the therapist to see an overview of the sessions. The timer feature was also found to be very useful. It gave information when one stage was about to finish so the

therapist can plan the next action to be carried out in the therapy session. The print function was also discovered to be very useful. Most therapists supported their answer by stating that the report will be used in the future, to know the overview what the patient did during the sessions and also to give feedback to the client.

Four therapists stated their preferences to fill in the patient and session information in the same user interface with the world control. One therapist stated that it actually does not matter, as long as this feature exists. Five therapists agreed that the same form for patient and session information entry was easier to use than the separated ones. Therapists also liked the flight control subjectively. It helped the therapist in planning a session.

Three therapists liked the idea of given restrictions in controlling the VE, but two of them stated these limitations were very annoying and did not give them enough freedom, especially in controlling the voice announcements.

All real therapists agreed the new improved system was easier to use than the old system. One therapist said that the new system was more difficult to learn due to the more complicated features, but it has a structured user interface that makes it convenient to use. Three therapists stated the new system is easier to learn and one therapist said it did not matter. All therapists also agreed that they liked all the added features. Most of them like the lightning and thunder, because they gave a surprise feedback. The most useful feature of all was the flight view that combined options from flight control and voice control. The therapists also liked real voice announcements from pilot and purser, and the sound of flap wings and landing gear. One therapist stated that the report feature would be very useful. All therapists preferred to use the new system to treat a patient who has fear of flying. They said that the new system was more organized than the old system. One therapist who gave initial requirements states that we had almost everything fulfilled, except for the amount of the avatars and the unreal look of the clouds. One therapist suggested that we should have separated approaching, touch down and taxiing stages during aircraft landing.

DISCUSSION

Overall, all subjects gave positive feedbacks to the improvement of the therapist's UI. From sixteen therapists, ten of them state their preferences in using the new system instead of the old system. One therapist preferred the old system to the new one, and five therapists did not give their preferences. Most of them preferred the new system to the old system because of the language used, more controllable features, its ease of use, easily learnable, and it provides clearer instructions. One therapist preferred the old system because it was less complicated due to less number of buttons that needed to be pressed to operate the system.

The possibility to compose scenarios and to simply run them afterwards was coined by one of the therapists as his suggestion. We referred to this function as autopilot. We thought about this function in the beginning [of what???], but from initial interviews, the therapist wanted to have complete control during therapy session. Thus, this feature was not implemented. Other useful suggestions were the

introduction of cabin sound (people talking, baby crying, etc.) and alert sound for alarm.

Some feedbacks were also gathered and therapists were asked to list three things they liked most and least in using the system. Lightning and thunder became favorite features in the new system, followed by feature of information overview during therapy sessions with linked option and limitation. The possibility to see what the patient's sees in VE by the therapist also evaluated very well. The system gave therapists a feeling of full control during the therapy session. The overall sound effects in the old system were louder than the new system. It can clearly be heard during landing stage. We think it would be nicer if the new system could use the same quality of sound as in the old system. Most of the therapist did not like the alarm reminder that was not functioning very well in the new system. It did not produce a reminder alert. The note feature was also not too useful either.

CONCLUSION

We can conclude that our formulated hypothesis for usability is accepted. It was significantly better than the old system, i.e. it increases the usability for the therapist.

VRET is slowly becoming the daily practice of therapists. During this transition, the usability issues play an important role in the acceptance of such an advanced technology. In the coming years, more therapists will work with this VRET "fear of flying" systems. It is hoped that the usability improvement of this and other systems could make their work much easier and could possibly increase the effectiveness of their therapy. However, to convince therapists to start using VRET systems an important issue is to show the profits on the level of the general health system and its costs. It is already shown that VRET for acrophobia works (Emmelkamp et al. 2002) and achieves the same results as treatment in vivo. This means that all "other" profits such as flexibility, better control of the conditions, increase the scale of use, etc. are to be exploited. But the health system including insurance companies must get interested and involved.

The new system including the new therapist's user interface as evaluated here offers a considerable better function for the therapist to treat fear of flying using VRET. For the usability engineering method, it is interesting to note that the usability profits of the new system show the same pattern for students as for real therapists. It is often hard to get a large number of specialists involved, and adding non-specialists can help to collect sufficient data on user behaviour (cf. user interfaces for astronauts in the space domain; Neerinx et al, 2001). By a better user interface for the therapist we think that the therapy for this kind of phobia can be done more efficient, more effective and with more satisfaction. The interface could be further improved by offering some support to the therapist in the form of a built-in agent advising the therapist on the next steps to take.

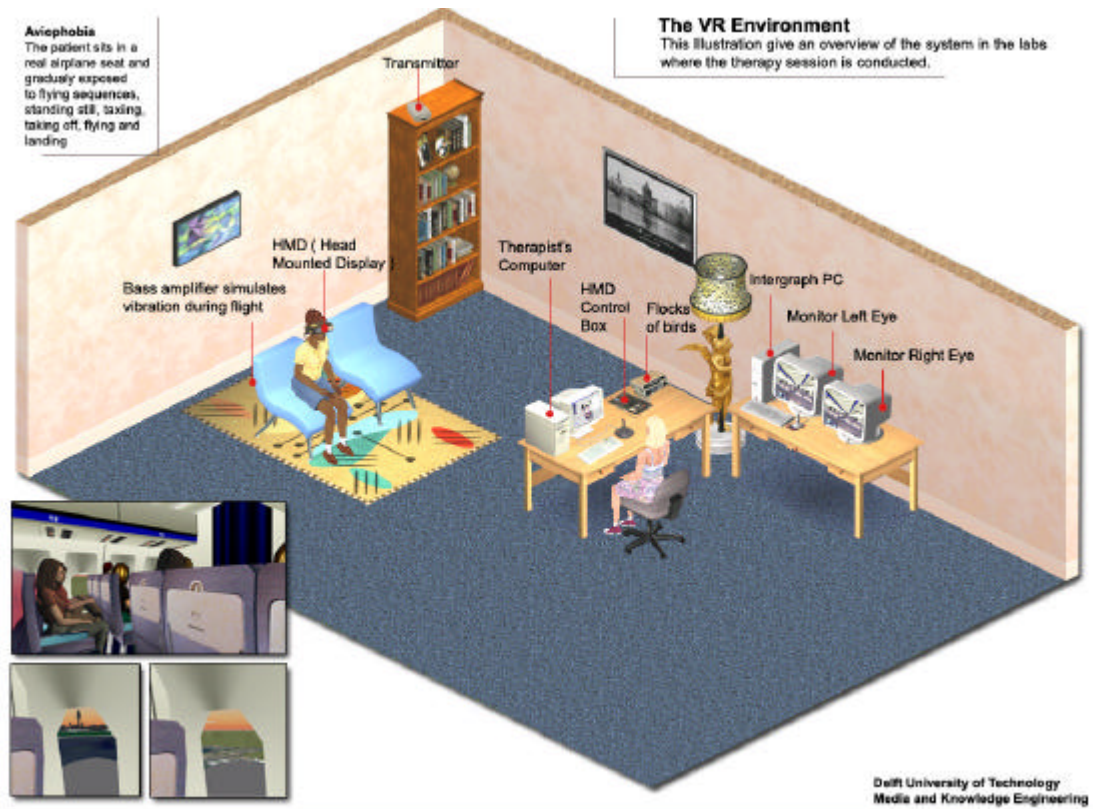
ACKNOWLEDGEMENTS

The authors like to thank Lucas van Gerwen from Stichting VALK in Leiden for his comments and for his support as director of a team of fear of flying therapists.

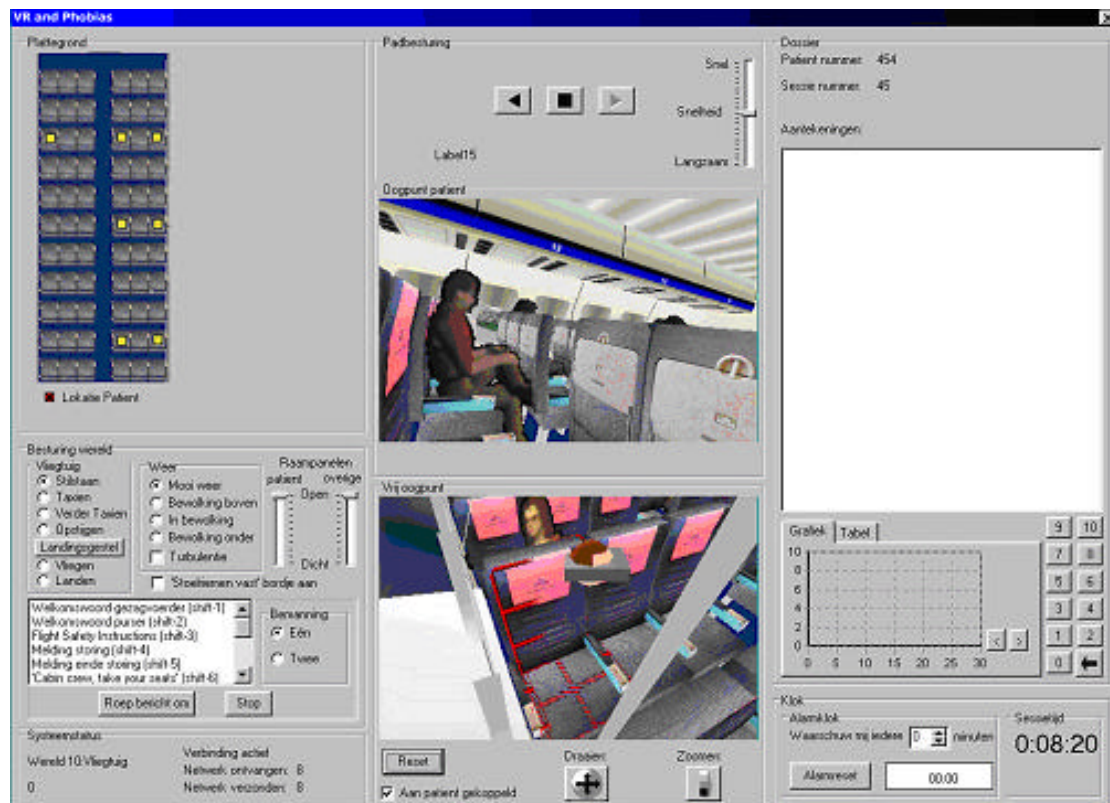
REFERENCES

- Emmelkamp, P.M.G.; M. Krijn; A.M. Hulsbosch; S. de Vries; M.J. Schuemie; and C.A.P.G. van der Mast. 2002. *Virtual Reality Treatment versus exposure in vivo: A Comparative Evaluation in Acrophobia Behaviour Research & Therapy*, Vol 40(5), pp.509-516.
- Neerinx, M.A.; M. Ruijsendaal; and M. Wolff. 2001. Usability Engineering Guide for Integrated Operation Support in Space Station Payloads, *International Journal of Cognitive Ergonomics*, 5(3), 187-198.
- Nielsen, Jacob. 1993. *Usability Engineering*. Boston: AP Professional.
- Rosson, Mary Beth; and John M Carroll. 2002. *Usability engineering scenario-based development of human-computer interaction*. 1st ed. Morgan Kaufmann Series in Interactive Technologies. San Fransisco: Academic Press.
- Schuemie, Martijn; and Charles van der Mast. 2001. *VR Testbed Configuration for Phobia Treatment Research*. Proceedings of the Euromedia 2001 Conference, April 18-20 2001, Valencia, Spain, pp.200-204.
- Schuemie, M.J.; C.A.P.G. van der Mast; M. Krijn; and P.M.G. Emmelkamp. 2002. *Exploratory Design and Evaluation of a User Interface for Virtual Reality Exposure Therapy*, in: J.D. Westwood, H.M. Hoffman, R.A. Robb, D. Stredney (Eds.), *Medicine Meets Virtual Reality 02/10*, IOS Press, pp.468-474
- Schuemie, Martijn. 2003. *Human-Computer Interaction and Presence in Virtual Reality Exposure Therapy*. Ph.D. Thesis, Delft University of Technology, available via <http://graphics.tudelft.nl/~vrphobia>.

Appendix A. The VR environment for treating phobias at Delft University of Technology



Appendix B. The old user interface for the therapist to control the sessions.



Appendix C. The new user interface for the therapist to control the sessions.

