Proceedings of Virtual Reality International Conference (VRIC 2011), 6-8 April 2011, Laval, France.
RICHIR Simon, SHIRAI Akihiko Editors. International conference organized by Laval Virtual.

Design and Evaluation of a Virtual Environment for the Treatment of Anger

BRINKMAN, Willem-Paul[1], HATTANGADI, Neeti[1], MEZIANE, Zakaria[1], PUL, Peter[1]

[1]Delft University of Technology, The Netherlands
w.p.brinkman@tudelft.nl

Abstract—**Cognitive-behavioural therapy is often used for anger treatment. An important element of this therapy is exposure to anger evoking stimuli. In this paper virtual reality is put forward as a technology that can effectively create these stimuli by exposing patients to social scenes that include anger stressors such as aggressive dialogues with virtual characters or arousing surrounding with loud music or flashing light. Applying a situated cognitive engineering approach a prototype system was developed which allowed a therapist to control these stressors. To evaluate the prototype an experiment was conducted in which participants, 18 non-patients and 2 patients, were exposed in a virtual environment to three types of social scenes: (1) a passive dialogue, (2) an aggressive dialogue, and (3) an aggressive dialogue with arousing surrounding. Results showed that these conditions had a significant effect on participants' galvanic skin response and the type of verbal reply towards the avatar. This effect was significant larger for the two patients than the non-patient group. In addition, evaluation of the therapist user interface suggested that most interaction components were relatively easy to use.**

Keywords: anger; anger management; virtual reality exposure therapy; design; patient; therapist.

I. INTRODUCTION

Anger can be both a healthy and an unhealthy emotion. It can be a driving force for example to stand up for someone's rights. However, it can also take the form of tantrums, rages and hate, resulting in aggressive and violent behaviour even as a response to minor or irrelevant provocations. Anger is often identified as one of the basic human emotions [1, 2]. Although the causes that might lead to this emotion can vary, trait anger, anxiety, depression, stress, and exposure to violence have been identified as anger predictors with moderate to substantial effect sizes [3]. As several authors [4, 5] have pointed out, anger disorders are currently not recognised by the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV-TR) [6]. Still difficulties with anger are observed in disorders mentioned by the DSM-IV-TR such as intermittent explosive disorder (IED), which is an impulse-control disorder which is characterized by discrete episodes of failure to resist aggressive impulses resulting in serious assault or destruction of property. Estimations [7] are that at some times during their life 7.3% of the US population might suffer from IED. The first anger attack is on average in early adolescence (14 years), with a mean of 43 lifetime attacks estimated to result in $1359 in property damage. Various psychological treatments are available for maladaptive anger, such as cognitive, cognitive behavioural therapy, exposure, psychodynamic, psycho-educational, relaxation-based, skill-based, stress inoculation, and multicomponent. On average these treatments are moderately successful at reducing anger problems [4]. Cognitive-behavioural therapy (CBT) is the most common approach to anger management, and its recipients are better off than 76% of untreated individuals with regard to anger reduction [8]. A specific example of CBT is the stress inoculation training (SIT) [9], which aims at the reduction and prevention of stress. Besides preparing individuals for stressful events such as military combat, medical surgery, and divorce, SIT has also been applied for the treatment of adolescents and adults with anger control problems [10]. SIT consists of three interlocking phases: (1) a conceptual educational phase which focuses on establishing a collaborative relationship with the patient and on helping them with understanding the nature and impact of their stress and coping resources; (2) a skill acquisition and skill consolidation phase which focuses on helping patients acquire coping skills and rehearse these; and (3) an application and follow-through phase which focuses on opportunities for patients apply these coping skills across increasing levels of stressors. Especially in this last phase patients rehearse their skills in vitro and gradually in vivo. A considerable amount of research has been conducted to study whether the often costly and difficult to organise in vivo exposure can be replaced by exposure in virtual reality (VR) for anxiety disorders such as claustrophobia, fear of driving, acrophobia, fear of flying, spider phobia, social phobia, panic disorder with agoraphobia, and posttraumatic stress disorder. Several meta-studies [11-13] on reported efficacy treatment studies, suggest that exposure in VR is as effective as exposure in vivo. For treatment of specific phobias, 76% of patients in a survey [14] indicated a preference for exposure in VR over exposure in vivo, and treatment refusal rate was far lower for exposure in VR (3%) than in

vivo (27%). Some promising research [15] has also been conducted in the use of VR for the treatment of anger. Anger arousal occurred when individuals were exposed in VR 360° panoramic video environment to anger provoking video. The exposure however is still limited in its social engagement. For example, for the treatment of social phobia, work [16, 17] has been done to engage patients into a natural dialogue with virtual characters, i.e. avatars, in the VR environment. The work reported in this paper, therefore, studies the effect of applying this technology for VR exposure as part of anger treatment, specifically the possibility of creating anger evoking stimuli. This research focuses on design and development of the technology. This is referred to as stage one in the collaborative cycles for the design and evaluation of technology for mental health interventions [18]. Stage two is the clinical evaluation which is outside the scope of the work presented here. In stage one, because of ethical considerations, access to patients is limited and the evaluation focus is on the usability of the technology. Because of its explorative nature a situated cognitive engineering [19] approach was followed, focusing both on the patient and therapist side of the system. After initial literature study and multidisciplinary meetings between software engineers and health care providers, a set of use scenarios were defined. Reviewing these scenarios led to requirements that were used as starting point of several prototype design iterations. The final prototype was evaluated in experiment with a group of non-patients and with two patients. The following sections will discuss these activities and its obtained insights into more detail. The paper ends with a discussion of possible future research.

II. DESIGN

A key factor in a successful design of a mental health care support system is the interplay between software engineers and health care providers, where the later also ensure the essential translation to the clinical practice. The system was specifically designed for the conditions that applied for De Fjord, a Dutch mental health clinic for adolescence. The clinic has several indoor and outdoor patients with anger problems. Both a therapist and the e-health manager of the clinic participated in the multidisciplinary team.

A. Use Scenarios Analysis

Several reports [20, 21] have demonstrated the usefulness of use scenarios analysis to establish requirements for mental health care support systems. Therefore three use scenarios were developed that addressed claims about the physical setup of the treatment session. Although no specific treatment protocol was defined, VR exposure could be seen as being an element of the application and follow-through phase in SIT, which often includes patients to imagine provocative situations [22, 23] or to role play them with the therapist [24]. In all three scenarios the patient is standing in front of a screen on which the VR environment is projected (Figure 1). In this environment patients talk with virtual humans, as they have been successfully used in several other domains [25].

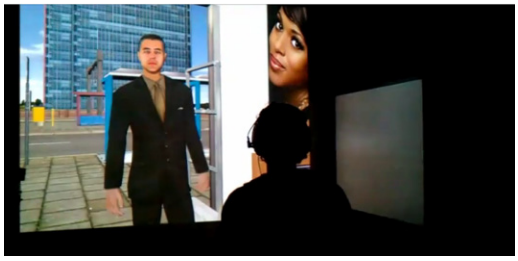

Figure 1: VR environment is projected on a life-size screen in front of the patient.

Figure 2: Scenario 1- left, therapist and patient separated during the VR exposure session; right, afterwards therapist and patient discuss the session.

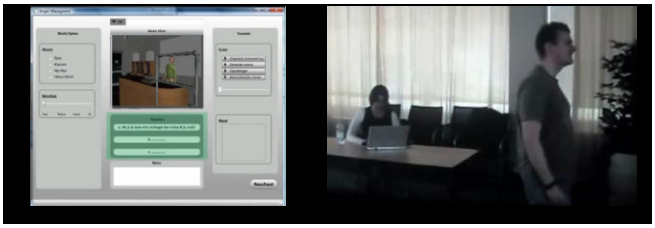

Figure 3: Scenario 2, left, avatar is directly controlled by therapist based on the observed behavior of the patient; right, therapist and patient are in the same room.

Figure 4: Scenario 3 - left, patient is received VR exposure in group setting; right, afterwards, group reflects on the session.

The first use scenario explored a setup where the therapist is not present during the exposure (Figure 2). The dialogue progresses in a linear fashion, whereby the avatar replies are not tailored to specific response of the patient. After the session, the therapist and the patient together reflect on the session. This scenario has cost efficiency advantages because of the reduced involvement of the therapist. The underlying claim is that the exposure does not require intervention by the therapist. Contra to this, the second scenario (Figure 3) gives therapists an active role in the exposure as they have to control the avatar, such as selecting verbal response and behaviours. As therapists are in the same room with the patient, they can tailor the dialogue to the (verbal) reactions of the patient. The third scenario (Figure 4) places the exposure in a group setting. Here other patients can also learn as they can observe how a dialogue progresses based on

reactions of their fellow patient. Afterwards, the group can also participate in the reflection of the exposure session. Each scenario formed the basis for a short film. The films were reviewed by the multidisciplinary team and resulted in a set of core functions (TABLE I), which were later modified and extended during the design of various prototypes.

TABLE I. CORE FUNCTIONS OF VR EXPOSURE SYSTEM FOR THE TREATMENT OF ANGER

No	Core Function
1	Therapist control avatar response
2	Multiple social scenes
3	Arousal enhancing environmental stressors
4	Recording behaviour and emotional state
5	Support for reflection

B. Prototypes

Insights obtained from the use scenario analysis were used as a starting point for three iterations of the design cycle in which the prototypes matured, starting with paper based prototype, followed by a low fidelity prototype, and concluding with a high fidelity prototype. All five identified core functions, as specified in TABLE I, were implemented in the final prototype. Figure 5 shows a screen shot of the user interface that the therapist used during the exposure session.

1) Therapist Control Avatar Response

To enhance their interaction in the social scene, patients should be able to freely talk with the avatars in the virtual environment. Although the avatar always starts the dialogue, the avatar should give appropriate replies to patients. Therefore, during session, the therapist can select verbal responses from a set of possible responses. If possible the verbal responses for both avatars and patients were labelled according to three types of response often used in assertiveness trainings [26]: (1) sub-assertive (passive) reactions, which try to avoid confrontation and do not offend a persons' rights; (2) assertive reactions, which affirms a person's right or point of view with respecting the another's rights; and (3) aggressive reaction which threatens the rights of others. For example, in a scene where a security guard stops the patient after the detection alarm goes off when the patient leaves the shop, the therapist could select from the following two verbal responses:

- <Assertive reaction> "Good afternoon sir, could I have a look into your bag please?"
- <Aggressive reaction> "Hey you! Give me your bag!"

The idea of using VR for assertiveness training is not new as it has also been suggested in helping adolescents to practice assertiveness skills to avoid or stop smoking [27].

2) Multiple Social Scenes

As multiple exposures are often needed, multiple social scenes are needed. Exposing a patient repeatedly to the same social scenes might not support patients to generalise their coping strategies across social situations. Although all scenes in the final prototype were situated in a clothes shop, the therapist could select from four social scenes (TABLE II). To set the setting, all scenes where introduced by a voice over of a storyteller.

3) Arousal Enhancing Environmental Stressors

The patients' coping resources might be influenced by non human stressors in the environment. Including these stressors in the exposure allows patient to rehearse their skills also in these situations. Therefore, the therapist had the option to select specific background music in the shop from a rage of themes: rock, hip-hop, heavy-metal, and classical music. In addition, the therapist could set the lighting in the shop from normal, very bright, to flashing (i.e. disco).

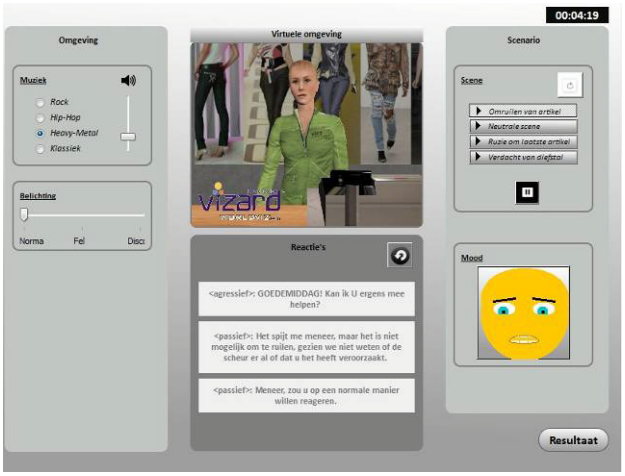

Figure 5: Therapist Control Center.

TABLE II. SOCIAL SCENES IN CLOTHES SHOP

No	Scene
1	*Neutral scene* where patient looks around in the shop and being greeted by the shop assistant and the security guard
2	*Changing a bought item*, where the patient negotiate with the shopping assistant on the return of an item
3	Argument with other customers about *the last item*, where the patient talks with customer and shopping assistant
4	*Suspect of stealing*, where after buying shoes the alarm goes off when leaving the shop and the patient is stop by security guard for security check

4) Recording Behaviour and Emotional State

During the exposure the patient might experience and display different emotions and (verbal) behaviours. To monitor the patient progress therapists might like to record these. In addition, this information could also be useful for the patient when discussing and reflecting upon the sessions. Although technologically labelling the behaviour of a patient by a therapist is possible, this would increase the therapist's workload. This is considered to be undesirable as it diverts attention from the treatment to the

management of the system [28]. Instead therefore an indirect approach was taken by recording the behaviour of the avatar. Besides giving an insight of the exposure, it could also give an insight into patient replies if therapists use a mirroring strategy in which the type of avatar reply matches the type of patient response, e.g. aggressive reaction as a response to a aggressive reaction, and non aggressive reaction (passive or assertive) as a response to a non aggressive reaction. The mirroring strategy seems reasonable to show the patient that a change in their aggressive behaviour would also result in less aggressive behaviour of other people. By using the affectbutton [29] the therapist was able to record the emotion state of the patient. This interaction component is an abstract face which emotional expression changes by moving the mouse cursor over it. By clicking on the face the emotion state is recording in pleasure (e.g. valence), arousal, and dominance (PAD) scores.

5) Support for Reflection

As mentioned before, after the session, the therapist and the patient might like to reflect on the session. As Figure 6 shows, the prototype presents in a table how the patient's emotion, the environmental stressors, and the avatars behaviour progressed over time. Although the emotion is recorded in PAD scores, the emotion is presented by an affect word (e.g. sleepy, frustrated, and angry) which position is most closely to the recorded PAD coordinates. Besides a textual overview, PAD scores can also be viewed in a graph (Figure 7).

III. EXPERIMENT

The system has two main users, the patient and the therapist, both with their own user interfaces. Both sides were evaluated. The focus on the patient side was to see if the virtual environment (VE) could evoke stress, and to study the level of presence and the experience of the dialogues. On the therapist side the focus was on the usability of the systems. The design and evaluation of technology for mental health support systems often follow two stages [18]. In the first stage the focus is on development of the technology whereas the second stage focuses on the clinical evaluation. The evaluation presented here therefore is located in first stage.

A. Method

The patient side experiment was setup as a within-subject design, in which participants were exposed to three experimental conditions: (1) passive dialogue, (2) aggressive dialogue, and (3) aggressive dialogue with additional environmental stressors (i.e. loud music and flashing light). The difference between passive and aggressive dialogue was created by the initial replies of the avatar. The response of the patients determined the progress of the dialogue whereby the therapist applied a mirroring strategy. To control for potential order effects the order of the conditions was countered balanced. For each condition a participant would see another scene. The three scenes (changing a bought item, last item, and suspect of stealing) were roughly evenly distributed over the conditions and sequence order.

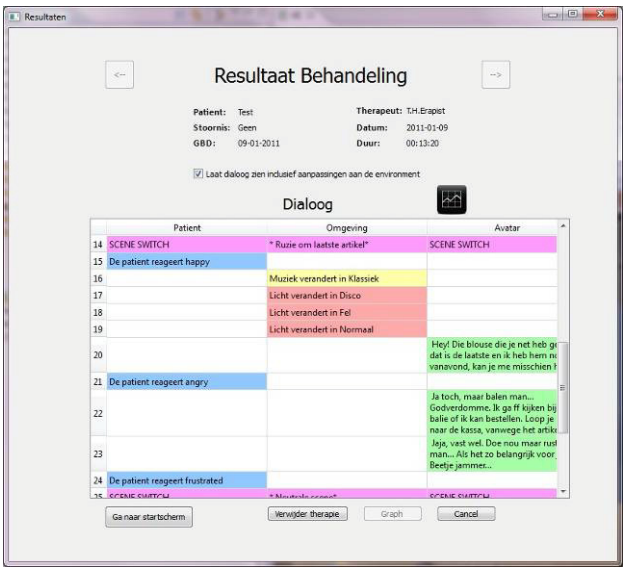

Figure 6: Result screen which shows in the table over time the affective state of the patient, environmental stressors, and the avatar responses.

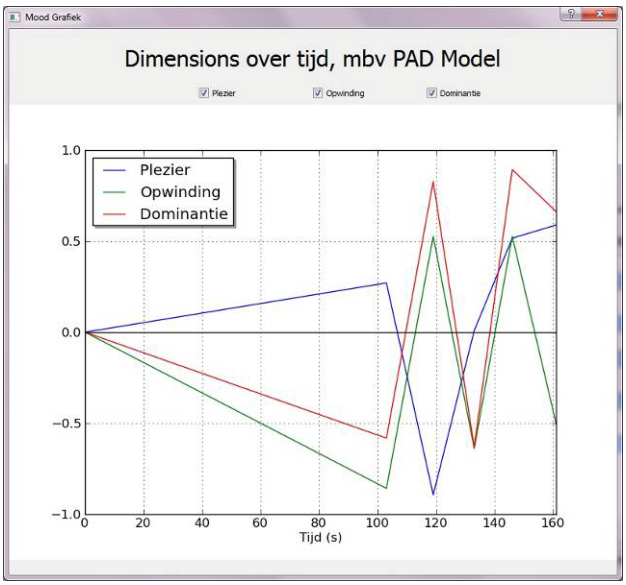

Figure 7: Affect graph based on Pleasure, Arousal, and Dominance model, progressing from neutral to sleepy, angry, sleepy, excited, and ending with relaxed.

1) Participants

The patient side was evaluated by a group of non-patients (N = 18) and two patients. The non-patient group consisted out of 14 males and 4 females with age ranging between 17 and 24 years (M = 21.9, SD = 1.75). These participants follow (pre-) university education and most with a technical focus. The two patients were patients of De Fjord clinic. The non-patient group also participated in the evaluation of the therapist side of the system.

2) Measures

For the evaluation of the patient side of the system, psychological measurement included the galvanic skin response (GSR) of the participants. This was measured with a standalone GSR Biofeedback Meter from Psychometric Research. Two electrodes were placed on two separate fingers of the participants.

The behaviour of the participants was recorded indirectly by recording the type of avatar response (passive, assertive, aggressive) selected by the experimenter. Subjective measure included:

- Simulation Sickness Questionnaire (SSQ) [30] was used to measure whether participants experienced symptoms that were associated with simulation sickness. The measure was administered before and after the experiment. The English version of the questionnaire was used with a Dutch translation.
- Reduced version of the Igroup Presence Questionnaire (IPQ) [31] was used to measure the overall experience level of presence (G1) and experienced realism in the virtual environments. The Dutch version of the inventory was used.
- Self-Assessment Manikin (SAM) [32] was used to measure experienced emotional state based on three PAD dimensions.
- The affectbutton was used to record observed emotions by the experimenter who acted as the therapist when controlling the exposure.
- Dialogue Experience Questionnaire (DEQ) [33] was administered to measure participants' experience of the dialogue and the avatars. The Dutch version of the inventory was used.

The therapist user interface was evaluated with the English component-based usability questionnaire (CBUQ) [34], in which participants were asked to evaluate: from start screen, the patient and therapy control; from the therapy control center, the world control, avatar reaction, scene control, and affectbutton; and the entire result screen.

3) Procedure

Participants of the non-patient group started the experiment by signing a consent form. After this they completed SAM, and SSQ. The experiment started with exposure in the neutral scene. This was followed by exposures in three conditions. During each exposure GSR responses were recorded every 10 seconds. After each of the three exposure sessions participants filled out SAM, DEQ and reduced IPQ, and once they had completed all sessions the participants filled out the SSQ again. During the exposure participants stood in a dark room and wore a headset. Two meters in front of them was a 3.5 by 2.5 meters screen on which the VE was projected.

The evaluation of the therapist side started with participants reading a small manual and then attempting to complete a task list in which they had to interact with all interaction components they later on evaluated with CBUQ. During this task one of the experimenters played the role of a patient.

Whereas the experiment with the non-patient group took place at the university, the experiment with the two patients took place at the clinic, and only involved the patient side evaluation. Patients were not asked to fill out any questionnaires, and GSR level was only recorded at the start and at the end of the session. The system was again controlled by the experimenter. During the session the therapist, a health care manager and the other patient was also present. The first patient (patient A) progressed through the experiment starting with a passive dialogue (last item), followed by an aggressive dialogue (suspect of stealing), and ending by aggressive dialogue with additional environmental stressors (changing a bought item). The second patient (patient B) started with passive dialogue (suspect of stealing), followed by passive dialogue with additional environmental stressors (changing a bought item), and ended with aggressive dialogue with additional environmental stressors (last item).

B. Results

1) Patient Side

The first analysis focused on the reported simulation sickness before and after the experiment. As patients had no control over their movement in the world, this could have caused simulation sickness. Simulation sickness could impact how participants would have experienced the virtual environment. To study the effect of the exposure a MANOVA with repeated measure was conducted with as dependent variables total SSQ score and three dimensions: nausea, oculomotor, and disorientation. The results shows an significant overall main effect ($F(3,14) = 4.01$, $p. = 0.030$) between pre and post measurement. This effect was also found back in unvariate analyses of total SSQ score ($F(1,16) = 12.71$, $p. = 0.003$) and in three dimensions nausea ($F(1,16) = 6.97$, $p. = 0.018$), oculomotor ($F(1,16) = 10.82$, $p. = 0.005$), and disorientation ($F(1,16) = 5.73$, $p. = 0.029$). Examining the means score (TABLE III) shows however that scores were lower after than before the VR exposure. This suggests that the system in general is not causing simulation sickness.

TABLE III. RESULTS SIMULATION SICKNESS SCORE BEFORE AND AFTER THE VR EXPOSURES.

	Before		After	
Dimension	***M***	***SD***	***M***	***SD***
Total	28.8	25.3	18.0	17.4
Nausea	20.8	21.7	12.9	15.1
Oculomotor	29.9	24.5	19.2	17.8
Disorientation	22.1	24.1	13.1	18.7

The second analysis was on the level of presence (IPQ). TABLE IV shows reported mean score of different experimental conditions. Two MANOVAs with repeated measures were conducted taking the general level of presence (G1) or experienced realism as the dependent variable and the experimental condition as independent variable. The analyses found no significant main effect for both general level of presence ($F(2,16) = 0.01$, $p. > 0.05$) or experienced realism ($F(2,16) = 2.63$, $p. > 0.05$). As no difference was found, a difference in quality of VE is unlikely to be an alternative factor that would otherwise

explain difference found between the experimental conditions.

TABLE IV. PRESENCE RESULTS FOR THREE EXPOSURE CONDITIONS.

	General		Realism	
Condition	***M*[a]**	***SD***	***M*[a]**	***SD***
Passive	3.50	1.65	2.14	0.97
Aggressive	3.56	1.62	2.18	1.02
Aggressive & Env.	3.50	1.69	1.90	0.76

a. Scale from 0 to 6

The experience dialogues were rated by participants on a scale from one to seven. TABLE V shows the mean score on the various sub-dimensions in the different conditions. Two MANOVAs with repeated measures were conducted with conditions as independent variable and the sub-dimension of a dimension as the dependent variables. No significant main effect ($F(8,10) = 0.45$, $p. > 0.05$) was found overall on the flow dimension nor in separate univariate analyses of the sub-dimensions. Also no significant main effect ($F(6,12) = 1.41$, $p. > 0.05$) was found overall on the interaction dimension. However, univariate analysis on reality dimension found a significant effect ($F(2, 34) = 3.89$, $p. = 0.030$). Post-hoc analysis revealed that the passive condition received a significant higher rating than the aggressive condition ($t(17) = 2.82$, $p. = 0.012$) and the aggressive condition with additional environmental stressors ($t(17) = 2.21$, $p. = 0.041$). This could mean that the passive dialogue was seen as more realistic either because the passive dialogue itself was realistic or that the aggressive element of two other dialogues was something these participants had not often experienced in real life and therefore unnatural for them. This second issue seems plausible as no difference was found between the conditions on realism IPQ dimension.

TABLE V. DIALOGUE EXPERIENCE RESULTS FOR THREE EXPOSURE CONDITIONS.

	Exposure Condition					
	***P*[a]**		***A*[b]**		***A & E*[c]**	
Dimension	***M***	***SD***	***M***	***SD***	***M***	***SD***
Flow						
dialogue speed	4.1	1.2	3.8	1.4	3.7	1.3
Interruption	5.2	1.3	5.4	1.2	5.5	0.9
correctness locally	4.3	1.1	4.2	1.3	3.9	1.0
correctness globally	4.7	1.3	5.3	1.2	4.9	1.2
Interaction						
Involvement	3.9	1.2	3.8	1.2	3.7	1.0
discussion satisfaction	3.9	0.8	3.9	0.9	3.7	0.9
Reality	4.3	1.1	3.6	1.1	3.7	1.1

a. Passive, b. Aggressive, c. Aggressive dialogue with additional environmental stressors

The SAM measurement directly addressed the self-reported emotional state of the participants. As can been seen in TABLE VI, means scores across the conditions were more roughly similar. A MANOVA with repeated measures on the three PAD dimensions also found no significant main effect ($F(6,12) = 0.37$, $p. > 0.05$).

The observed emotion recorded with the affectbutton, average for each session, is shown in TABLE VII. Again a MANOVA with repeated measures on the three observed PAD scores found no significant main effect ($F(6,11) = 0.50$, $p. > 0.05$) for the three conditions. The observed emotions of patient A and patient B was also recorded and their average for each condition can be seen in TABLE VIII. One-sample *t*-tests were conducted to examine whether the individual score of patient A or B significantly deviated from the mean of the non-patient group. As TABLE VIII shows, the patients' emotional state was less dominant which might suggest that these two patients felt less in control of the situation in the exposure. The two patients also seem to have some higher ratings on the pleasure dimension, which might suggest that they appeared to enjoy the exposure more than on average the non-patient group. The patients also mentioned that they like doing this exercise.

TABLE VI. SELF-REPORTED EMOTION RESULTS FOR THREE EXPOSURE CONDITIONS.

	Exposure Condition					
	***P*[a]**		***A*[b]**		***A & E*[c]**	
Dimension	***M***	***SD***	***M***	***SD***	***M***	***SD***
Pleasure	3.1	1.4	3.1	1.8	3.1	1.9
Arousal	6.0	1.8	6.4	2.1	6.1	1.8
Dominance	5.2	1.7	5.4	1.6	5.0	1.8

a. Passive, b. Aggressive, c. Aggressive dialogue with additional environmental stressors

TABLE VII. OBSERVED EMOTION RESULTS FOR THREE EXPOSURE CONDITIONS OF NON-PATIENT GROUP.

	Exposure Condition					
	***P*[a]**		***A*[b]**		***A & E*[c]**	
Dimension	***M***	***SD***	***M***	***SD***	***M***	***SD***
Pleasure	0.0	0.3	-0.1	0.3	-0.1	0.3
Arousal	-0.8	0.3	-0.8	0.3	-0.8	0.4
Dominance	0.2	0.2	0.2	0.1	0.2	0.2

a. Passive, b. Aggressive, c. Aggressive dialogue with additional environmental stressors

TABLE VIII. OBSERVED EMOTION RESULTS FOR THREE EXPOSURE CONDITIONS OF PATIENT A AND B.

	Exposure Condition					
	***P*[a]**		***A*[b]**		***A & E*[c]**	
Dimension	***A***	***B***	***A***	***B***	***A***	***B***
Pleasure	0.24*	-0.07	0.01		-0.20	0.11*
Arousal	-1.00*	-1.00	-1.00		-1.00	-1.00
Dominance	0.25	0.08*	0.04*		0.13*	-0.06*

a. Passive, b. Aggressive, c. Aggressive dialogue with additional environmental stressors,

* $p. < 0.05$

The GSR recordings gave an insight into the physiological response of the participants. As multiple recordings were made during the session, the median was taken to present the central tendency of GSR in an exposure session (δ). To correct for individual difference (α) this measurement was set against the measurement obtained in the neutral exposure session (β) that preceded the three exposure sessions (1).

$$\alpha = \frac{\delta - \beta}{\beta} \quad (1)$$

TABLE IX shows the mean scores of the relative GSR results of the non-patient group for the three conditions. A MANOVA with repeated measures found a significant main effect ($F(2,16) = 7.35$, $p. = 0.005$) for the conditions on these relative GSR scores. Post-hoc analysis only found that participants sweat significantly ($t(17) = 2.54$, $p. = 0.021$) more in the aggressive condition than the passive condition, which suggest a higher arousal level. The relative GSR scores at the end of each session for patient A and B can also be seen in TABLE X. One-sample *t*-tests found that both these two patients sweat significantly more than on average the non-patient group in the aggressive dialogue with additional environmental stressors. This suggests that whereas the additional environmental stressors had no significant effect on the non-patient group it did have an effect on these two patients.

TABLE IX. RESULTS GSR MEASUREMENTS FOR THREE EXPOSURE CONDITIONS.

	Non-patient		Patient	
Condition	***M***	***SD***	**A**	**B**
Passive	0.08	0.23	0.03	0.10
Aggressive	0.15	0.16	0.14	
Aggressive & Env.	0.13	0.14	0.28*	0.22*

* $p. < 0.05$

To analyse the behaviour of the participants, the selected avatar responses were analysed. First, the responses were quantified on a single dimension. Sub-assertive (passive) reactions were scored as 1, assertive reactions as 0, and aggressive reactions as -1. Next, for each participant the mean score was calculated in each session, excluding the opening response of the avatar. As the therapist was using a mirroring strategy, this mean score established an index for the responses of the participants. A MANOVA with repeated measures found a significant main effect ($F(2,15) = 10.71$, $p. = 0.001$) for the conditions on the behaviour index. Post-hoc analysis revealed that the behaviour score in passive condition was significantly higher than the behaviour score in the aggressive dialogue with additional environmental stressors condition ($t(16) = 4.77$, $p. < 0.001$). Looking at TABLE IX, it seems that in the passive dialogue participants' responses where more assertive towards sub-assertive, whereas in the other two aggressive dialogue conditions responses were more assertive towards aggressive responses. The behaviour index scores of the two patients seem more polarised (TABLE IX). One-sample *t*-tests results suggest that these two patients give a significantly more passive reply in the passive condition than the non-patient group. Furthermore patient A also gave a significantly more aggressive reply in aggressive condition than on average the non-patient group. For the last condition the results of the patients show two different trends. Whereas patient A gave more assertive/sub-assertive replies, patient B gave more aggressive replies compared to the non-patient group.

2) Therapist Side

The non-patients group also evaluated the therapist side of the system with CBUQ. The obtained mean score for each component was compared with the norm value of 5.29 [34]. With the exception of affectbutton component, all components had score significantly higher than the norm value. This means that the usability of these components was more comparable to the easy to use components than the difficult to use components from the norm set.

TABLE X. RESULTS BEHAVIOUR INDEX FOR THREE EXPOSURE CONDITIONS.

	Non-patient		Patient	
Condition	***M***	***SD***	**A**	**B**
Passive	0.39	0.55	1.00*	1.00*
Aggressive	-0.06	0.75	-1.00*	
Aggressive & Env.	-0.34	0.33	0.00*	-1.00*

* $p. < 0.05$

TABLE XI. RESULTS USABILITY EVALUATION INTERACTION COMPONENTS OF THERAPIST USER INTERFACE.

Component	M^a	***SD***	***t***	***p***
Start sceen				
patient control	6.6	0.56	9.54	< 0.001
therapist control	6.4	0.71	6.59	< 0.001
Therapy control center				
world control	6.3	0.80	5.16	< 0.001
avatar reaction	6.0	0.97	3.10	0.007
scene control	6.5	0.48	10.86	< 0.001
affectbutton	5.0	1.47	-0.912	> 0.05
Result screen	6.1	0.65	5.52	< 0.001

a. Scale from 1 to 7

IV. CONCLUSION AND FINAL REMARKS

The main conclusion that might be drawn from the experiment is that the system allows therapist to engage individuals in potential aggressive dialogues which increases their physiological arousal and evoke more aggressive replies. The system therefore seems to meet an essential requirement for exposing patients to potential aggressive social situations. The results of the two patients seem to support this finding. More than the non-patient group, they also seem to be affected by the additional

environmental stressors, and exhibited an emotional state of being less in control of the exposure situation.

Although promising, the experiment also had a number of limitations that should be identified. First, only two patients were involved and their data was collected in another physically and social environment, i.e. the clinic with actual health care providers present, than the non-patient group, i.e. at university in the lab with only the experimenters present. Because of practical reason no questionnaire data was collected from the patients. Still other data provides some explorative insight into how patients might response to these systems. Another issue is the individual variations between the three experimenters who controlled the therapist side of the system which might also have caused variance in the exposure and the report of the observed emotion. As the same experimenter acted as the therapist throughout all the conditions for a specific participant, it seems unlikely to have caused a systematic bias because of the within-subjects design of the experiment. Another contribution made by the study is the design of the system and the identified core functions. Usability of the various component of the therapist user interface was also perceived well by the participants. Although a promising indication, this should not directly be generalised to actual therapists.

Future works should also address how this system could be used effectively in the treatment of anger. As multiple exposures might be needed, the system might need to be extended to include more dialogues and also more physical settings. To reduce the workload of the therapist, work could also focus on using automatic free speech interaction which has also been explored in VR systems for treatment of social phobic patients [33]. In that context and in the context of anger treatment, engaging patients in natural dialogue seems to give the therapist the ability to provide the exposure in a controlled environment. Adding additional environmental stressors such as loud music and flashing light could give therapists the ability to put the patient in an emotional state necessary to evoke an aggressive response towards the avatar as a reaction to its behaviour which the patient might otherwise easily resist without the impulse to react aggressively.

ACKNOWLEDGMENT

We thank the health care providers at De Fjord for their support and involvement in this project.

REFERENCES

[1] S. C. Widen and J. A. Russell, "Young children's understanding of others' emotion," in *Handbook of emotions*, 3rd ed, M. Lewis, J. M. Haviland-Jones, and L. F. Barrett, Eds. New York: The Guilford press, 2008, pp. 348-363.

[2] P. Ekman, *Emotions revealed*. New York: Henry Holt and company, 2003.

[3] N. E. Mahon, A. Yarcheski, T. J. Yarcheski, and M. M. Hanks, "A Meta-Analytic Study of Predictors of Anger in Adolescents," *Nursing Research,* vol. 59, pp. 178-184, May-Jun 2010.

[4] M. Saini, "A Meta-analysis of the Psychological Treatment of Anger: Developing Guidelines for Evidence-Based Practice," *Journal of the American Academy of Psychiatry and the Law,* vol. 37, pp. 473-488, 2009.

[5] B. O. Olatunji and J. M. Lohr, "Nonspecific factors and the efficacy of psychosocial treatment of anger," *Scientific review of mental health practice,* vol. 3, pp. 3-18, 2005.

[6] American Psychiatric Association., *Diagnostic criteria from DSM-IV-TR*. Washington, D.C.: American Psychiatric Association, 2000.

[7] R. C. Kessler, E. F. Coccaro, M. Fava, S. Jaeger, R. Jin, and E. Walters, "The prevalence and correlates of DSM-IV intermittent explosive disorder in the national comorbidity survey replication," *Archives of general psychiatry,* vol. 63, pp. 669-678, 2006.

[8] R. Beck and E. Fernandez, "Cognitive-behavioral therapy in the treatment of anger: a meta-analysis," *Cogntive therapy and research,* vol. 22, pp. 63-74, 1998.

[9] D. Meichenbaum, *Stress inoculation training*. New York: Pergamon Press, 1985.

[10] D. Meichenbaum, "Stress inoculation training: a preventative and treatment approach," in *Principles and practice of stress management*, 3rd ed, P. M. Lehrer, R. L. Woolfolk, and W. S. Sime, Eds.: Guilford Press, 2007, pp. 497 - 516.

[11] M. B. Powers and P. M. Emmelkamp, "Virtual reality exposure therapy for anxiety disorders: A meta-analysis," *J Anxiety Disord,* vol. 22, pp. 561-9, 2008.

[12] L. Gregg and N. Tarrier, "Virtual reality in mental health : a review of the literature," *Soc Psychiatry Psychiatr Epidemiol,* vol. 42, pp. 343-54, May 2007.

[13] T. D. Parsons and A. A. Rizzo, "Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: a meta-analysis," *J Behav Ther Exp Psychiatry,* vol. 39, pp. 250-61, Sep 2008.

[14] A. Garcia-Palacios, C. Botella, H. Hoffman, and S. Fabregat, "Comparing acceptance and refusal rates of virtual reality exposure vs. in vivo exposure by patients with specific phobias," *Cyberpsychol Behav,* vol. 10, pp. 722-4, Oct 2007.

[15] S. D. Miyahira, R. A. Folen, M. Stetz, A. Rizzo, and M. M. Kawasaki, "Use of immersive virtual reality for treating anger," in *Annual review of cybertherapy and telemedicine 2010*, B. K. Wiederhold, G. Riva, and S. I. Kim, Eds. Amsterdam: IOS Press, 2010, pp. 82-86.

[16] W.-P. Brinkman, C. A. P. G. van der Mast, and D. de Vliegher, "Virtual reality exposure therapy for social phobia: A pilot study in evoking fear in a virtual world," in *HCI2008 workshop - HCI for technology enhanced learning*: Mediamatica, 2008, pp. 85 - 88.

[17] N. ter Heijden, C. Qu, P. Wiggers, and W. P. Brinkman, "Developing a dialogue editor to script interaction between virtual characters and social phobic patients," in *ECCE2010 Workshop on cognitive engineering for technology in mental health care and rehabilitation* Delft, The Netherlands: Mediamatica, Delft University of Technology, 2010, pp. 111-123.

[18] D. Coyle, G. Doherty, M. Matthews, and J. Sharry, "Computer in talk-based mental health interventions," *Interacting with Computers,* vol. 19, pp. 545-562, 2007.

[19] M. A. Neerincx and J. Lindenberg, "Situated cognitive engineering for complex task environments," in *Naturalistic decision making and macrocognition*, J. M. C. Schraagen, L. Militello, T. Ormerod, and R. Lipshitz, Eds. Aldershot, UK: Ashgate, 2008, pp. 373 - 390.

[20] C. Paping, W.-P. Brinkman, and C. van der Mast, "An explorative study into tele-delivered multi-patient virtual reality exposure therapy system," in *Coping with posttraumatic stress disorder in returning troops: Wounds of War II*, K. Wiederhold, Ed.: IOS press, 2010, pp. 203-219.

[21] M. van den Steen, W. P. Brinkman, E. Vermetten, and M. Neerincx, "Multi-modal memory restructuring for patients suffering from combat-related PTSD: a pilot study," in *Annual review of cybertherapy and telemadicine 2010*, B. K. Wiederhold, G. Riva, and S. I. Kim, Eds. Amsterdam: IOS press, 2010, pp. 208 - 213.

[22] P. L. Timmons, M. E. Oehlert, S. W. Sumerall, C. W. Timmons, and S. B. Borgers, "Stress inoculation training for maladaptive anger: Comparison of group counseling versus computer guidance," *Computers in Human Behavior,* vol. 13, pp. 51-64, Jan 1997.

[23] J. L. Bistline and F. P. Frieden, "Anger Control - a Case-Study of a Stress Inoculation Treatment for a Chronic Aggressive Patient," *Cognitive Therapy and Research,* vol. 8, pp. 551-555, 1984.

[24] K. J. Schlichter and J. J. Horan, "Effects of stress inoculation on anger and aggression management skills of institutionalized juvenile delinquents," *Cognitive Therapy and Research,* vol. 5, pp. 359 - 365, 1981.

[25] N. Magnenat-Thalmann and D. Thalmann, *Handbook of virtual humans*. Chichester, England ; Hoboken, NJ: Wiley, 2004.

[26] S. Hasan, "A tool to teach communication skills to pharmacy students," *American journal of pharmaceutical education,* vol. 72, 2008.

[27] K. Nemire, J. Beil, and R. W. Swan, "Virtual world for helping teens practice assertiveness skills," in *Steroscopic displays and virtual reality systems VI*, J. O. Merritt, M. T. Bolas, and S. S. Fisher, Eds., 1999, pp. 404-415.

[28] W.-P. Brinkman, C. van der Mast, G. Sandino, L. T. Gunawan, and P. Emmelkamp, "The therapist user interface of a virtual reality exposure therapy system in the treatment of fear of flying," *Interacting with Computers* vol. 22, pp. 299-310, 2010.

[29] J. Broekens and W. P. Brinkman, "Affectbutton: towards a standard for dynamic affective user feedback," in *Affective computing and intelligent interaction (ACII)*, 2009, pp. 1-8.

[30] R. S. Kennedy and E. L. Norman, "Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness," *The international journal of aviation psychology,* vol. 3, pp. 203-220, 1993.

[31] T. Schubert, F. Friedman, and H. Regenbrecht, "The experience of presence: factor analytic insights.," *Presence,* vol. 10, pp. 266-281, 2001.

[32] M. M. Bradley and P. J. Lang, "Measuring emotion: The self-assessment manikin and the semantic differential," *Journal of behavior therapy and experimental psychiatry,* vol. 25, pp. 49-59, 2002.

[33] N. ter Heijden and W. P. Brinkman, "Design and evaluation of a virtual reality exposure therapy system with automatic free speech interaction," *Journal of CyberTherapy and Rehabilitation,* in press.

[34] W.-P. Brinkman, R. Haakma, and D. G. Bouwhuis, "Theoretical foundation and validity of a component-based usability questionnaire," *Behaviour and Information Technology,* vol. 28, pp. 121 - 137, 2009.