

How technology influences the therapeutic process: A comparative field evaluation of Augmented Reality and In Vivo Exposure Therapy for phobia of small animals

Maja Wrzesien¹, Jean-Marie Burkhardt², Mariano Alcañiz^{1,3} Cristina Botella^{3,4}

¹ Instituto Interuniversitario de Investigación en Bioingeniería y Tecnología Orientada al Ser Humano, Universidad Politécnica de Valencia, Camino de Vera s/n, 46022 Valencia, Spain

² Paris Descartes University, LATI, 45 rue des Saints-Pères 75270 Paris cedex 06, France

³ CIBER, Fisiopatología Obesidad y Nutrición, CB06/03 Instituto de Salud Carlos III, Spain

⁴ Departamento de Psicología Basica y Psicobiología Universidad Jaume I, Castellón, Spain

Abstract. In Vivo Exposure Therapy (IVET) has been a recommended protocol for the treatment of specific phobias. More recently, several studies have suggested that Augmented Reality Exposure Therapy (ARET) is a potentially effective technology in this field. The objective of this paper is to report the preliminary results of a comparative analysis of ARET and IVET applied to the treatment of phobia to small animals. To analyze participants' activity, we have adopted a multidisciplinary and mixed perspective based on clinical and user-centered approaches. This pilot results show that ARET and IVET are both clinically effective. Both therapies produce a significant reduction in the clinical outcome measures and allow the clients to interact with a real phobic stimulus after the therapeutic session. The results also show some main differences between technology-mediated therapy and traditional non-mediated therapy. We discuss these results in terms of future design and evaluation guidelines for Mental Health technologies.

Keywords: Mental health, augmented reality, field evaluation.

1 Introduction

It is increasingly recognized that innovative technologies have strong potential in the Mental Health (MH) field [1]. New technologies such as Virtual Reality (VR) or Augmented Reality (AR) can provide therapists with a wide range of MH services and functions to support their therapy and assessment activities. Also, VR and AR allow clients¹ to have easier access to MH services and enhance their engagement in the treatments [2] due to the strong representational and immersion capability of these technologies. However, the way a specific MH system is designed may induce use-related problems. For example, a weak design of Human-Computer Interactions (HCI) may lead to a significant increase in the therapists' or clients' workload during

¹ The word client is usually used in Mental Health Care to describe a person suffering from mental illness

the clinical intervention. The task of using the system itself may also distract the clients as well as the therapist from the therapeutic session. All of these aspects have stimulated researchers to place more importance on the design guidelines and evaluation techniques for MH technologies (e.g. [1], [2], [3]).

A critical but still rather unexplored issue consists in how these technologies actually support their intended users in their respective (but closely related) activities; how they are actually used; and how they modify the therapeutic process. Among studies of mental health VR/AR systems, most have concentrated on a “client perspective” (e.g. [4], [5]) with a focus on measuring client outcomes as a function of the VR/AR systems and/or therapy orientations. On the other hand, of the few studies in HCI field regarding mental health VR systems (e.g. [6], [7]), most have concentrated on a “therapist perspective” with a focus on measuring or observing the activity of the therapists with the system. Both perspectives are limited in the sense that they consider only a single user at a time, instead of looking at both the therapist and the client working together in a virtual and/or real environment. Therefore, there is a need to propose the “mixed perspective” that should take into account the two actors of the therapeutic process (client and therapist) as well as the technology and the environment with virtual and/or real artifacts. Moreover, to the authors’ knowledge, no studies regarding HCI issues have yet been published in the field of mental health AR technologies.

The aim of this paper is to present a multidisciplinary (i.e., HCI and clinical) and mixed approach (i.e., one that takes into account the two actors of the therapeutic process, the technology, and the collaborative processes that take place among them in a virtual and/or real environment) in a case study of MH technology evaluation. More specifically, this paper reports pilot data on a clinical setting evaluation of Augmented Reality Exposure Therapy (ARET) and In Vivo Exposure Therapy (IVET) for small animal phobias (spiders and cockroaches). The paper is organized as follows. Section 2 presents some main issues regarding In Vivo, Virtual Reality, and Augmented Reality Exposure Therapies applied to the treatment of small animal phobias. Section 3 presents our framework to address the evaluation. Section 4 presents the methodology undertaken in this evaluation. The remainder of the sections present the results and conclusions with a number of design implications for the MH technologies that are drawn from the analysis.

2 In vivo, virtual reality, and augmented reality exposure therapy in the treatment of specific phobias of small animals

Specific phobias all share a common pattern of “marked and persistent fear of clearly discernible circumscribed objects or situations” [8]. Reports on Mental Health present specific phobias as one of the most common single mental disorders [9]. In Vivo Exposure Therapy (IVET) is considered to be standard therapy for the treatment of specific phobias [10]. During this type of therapy, the clients are exposed to real (live) phobic objects or situations, they confront real spiders in arachnophobia or cockroaches in phobia to cockroaches. The effectiveness of IVET has been demonstrated by numerous researchers; however, this treatment has some drawbacks.

First, the therapist is not in full control of the real phobic object or the situation. Thus, it is difficult for the client to have a personalized, graduated exposure in accordance with specific fears (i.e., the only way to present cockroaches or spiders without any movement is to kill them). Second, arranging and organizing the exposure can be time-consuming and can create logistics problems (e.g. feeding the animals; finding a place for the terrarium; etc.). Third, the treatment is associated with a high dropout rate and low level of acceptance [11].

Virtual Reality Exposure Therapy (VRET) might be an interesting option to consider when trying to motivate phobia sufferers to treat their disorders. Indeed, besides having the same effectiveness as traditional therapy [12], VRET resolves some of the problems associated to IVET [4]. First of all, VRET allows therapists to precisely control the phobic stimulus (i.e., the controlled virtual cockroach or spider can stay immobile, can be moved in different directions on a smaller or wider scale, can change size, and can be multiplied as many times as the client and therapist desire). Second, the virtual animals do not require anything to keep them alive; a simple click of the computer button is enough to make them appear. Third, VRET can recreate environments and situations that would be difficult to arrange in traditional treatments such as IVET.

Augmented Reality Exposure Therapy (ARET) has the same advantages as VRET, but it also has some additional characteristics that might be appealing to both clients and therapists. The ARET system allows clients to perceive the real environment and their body with virtual objects (i.e., augmented reality). According to Botella et al. [4], in the case of small animal phobias, this has two great advantages over the VRET applications. First, the real environment does not have to be modeled; therefore, the costs of programming and modeling the application are lower (limiting the modeling and programming to some specific objects such as cockroaches or spiders). Second, perceiving the virtual object in the real environment may have great importance for a better sense of presence and reality judgment, which are recognized as key aspects in this field [13].

ARET has not yet been thoroughly explored, especially regarding HCI issues. Therefore, the aim of this paper is to fill in some of the gaps in this particular research area and to propose a multidisciplinary and mixed perspective approach for evaluating this MH technology under real world conditions.

3 A multidisciplinary and mixed perspective approach

The goal of the Augmented Reality Exposure Therapy system is to allow both the therapist and the client to collaborate so that the client confronts the phobic stimulus and interacts with it at the lowest possible level of discomfort. In this paper, we are interested in studying two unexplored issues associated to this perspective. First, the design of the system should support the therapeutic relationship between the client and the therapist, which is defined as the therapeutic alliance. In fact, a different interpretation of the technology-mediated therapeutic process is needed. The dynamics of the therapeutic process has until now been analyzed from a purely clinical point of view. This is understandable when the therapeutic process

corresponds only to face-to-face traditional client-therapist interactions. However, with the introduction of new technologies to the therapist's office, the dynamics of the client-therapist interactions may change. Therefore, this new technology-mediated context should be addressed in terms of collaboration. Indeed, the relationship between the client and the therapist, which is defined in terms of the therapeutic alliance concept, is expressed as "(...) the quality and strength of the collaborative relationship between client and therapist (...)" [14]. The quality of this relationship has been shown to contribute to 30% of the positive changes in the client [15]; therefore its quality is of particularly great interest. This point of view is shared by different authors. Doherty et al. [1] proposed using the therapeutic alliance measure between client and therapist in order to gain insight into effects of technology on the therapeutic relationship. Meyerbröker and Emmelkamp [16] analyzed the relationship between the therapeutic alliance and the outcome of the therapy with Virtual Reality Exposure Therapy for specific phobias (i.e., acrophobia and fear of flying).

The second unexplored issue corresponds to the adaptation of the ARET system to face-to-face psychotherapy in an environment mixing both real and virtual objects. Previous HCI research studies in the MH field have not explicitly exploited the "mixed perspective" in order to account for the complex interactions between the therapist, the client, the technology, and the real and/or virtual environments in which they take place. Therefore, a systemic analysis (i.e., mixed) framework is needed in order to fully understand this complex therapeutic context. Recently, Nardi and Kaptelinin [17] showed the benefits of systemic analysis such as Activity Theory in the design of different applications; and Hollan and Hutchins [18] demonstrated the benefits of the Distributed Cognition framework in the design and evaluation of different Augmented Environments. Wrzesien et al. [19] applied a similar approach to analyze Augmented Reality Exposure Therapy. Thus, in our opinion, the application of a distributed theoretical framework has the following advantages. First, it takes into account all possible components of the unit of analysis (i.e., client, therapist, technology, environment, and the interactions among them). Second, the analysis can take place in real world conditions by taking into account day-to-day clinical practice and making the evaluation ecologically valid (i.e. performed in the real-life situation). Finally, the technology is considered to be an equally important component of the unit of analysis, which may positively and/or negatively influence other components.

In summary, the therapeutic process mediated by technology or non-mediated by technology should be analyzed in the following terms. First, the HCI issues and clinical issues should be combined into a multidisciplinary approach. Second, the mixed perspective analysis of interactions between the different poles involved in the therapeutic activity should be proposed. Therefore, the results in the following sections will reflect this double multidisciplinary and mixed approach.

4. Methodology

4.1 Research design

The study compares two different types of therapeutic processes: the traditional therapeutic process (IVET) and the technology-mediated therapeutic process (ARET).

These two therapeutic processes included identical therapeutic objectives and clinical protocols and were located in the same place. In order to reduce the impact of the individual therapeutic style on the development of the client-therapist relationship, all of the therapists treated clients in both groups. Thus, any differences in clients' clinical outcomes, clients' perceived therapeutic relationship, and therapeutic activities should be attributed to the therapy type factor (technology-mediated vs non-mediated).

4.2 Participants

The participants for this study (clients) were selected according to the DSM-IV [8] criteria for a specific phobia to small animals. In total, 12 clients participated in this study (eleven women and one man, $M=28.54$ years old; $SD=7.92$). Each client was randomly assigned to one of two groups (the ARET group or the IVET group).

Three therapists working in the clinic participated in this study. They all had a minimum of one year of experience in the therapeutic field; however, two of them were novices in the use of the ARET system.

4.3 Materials

All the therapeutic sessions followed the same “one-session treatment” protocol [20]. The protocol involves the use of intensive exposure carried out in one session of a maximum of three hours. The protocol is composed of four parts: (a) exposure; (b) modeling (by demonstrating the interaction with the phobic stimulus by the therapist followed, if possible, by the client); (c) cognitive restructuring; and (d) reinforcement. The main goal of the therapeutic session is to allow the client to confront the phobic stimulus and interact with it at the lowest possible level of discomfort.

a) In Vivo Exposure Therapy

The IVET (see Figure 1a) corresponds to the direct confrontation of a real feared stimulus (i.e., cockroaches or spiders). This type of therapeutic activity involves both the therapist and the client interacting with a real cockroach or spider in order to expose the client to his/her phobic stimulus.

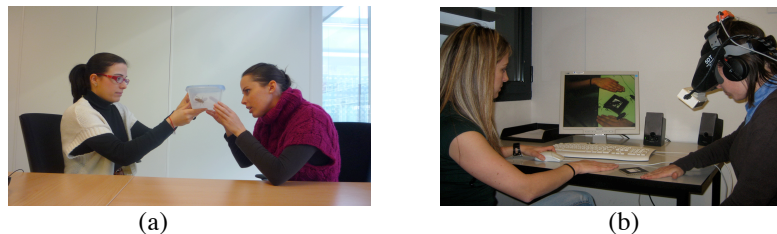


Fig. 1. In Vivo Exposure Therapy (Figure 1a), and Augmented Reality Exposure Therapy (Figure 1b). Therapists (on the left) and clients (on the right) interact with the phobic stimuli

b) Augmented Reality Exposure Therapy

ARET (see Figure 1b) corresponds to the direct confrontation of a virtual feared stimulus in the real environment (i.e., augmented reality). This type of therapeutic activity involves both the therapist and the client interacting with the virtual cockroach or spider in order to expose the client to his/her phobic stimulus.

The interface that controls the ARET system was only available to the therapist and corresponded to a keyboard. The system allows the therapist to control the phobic stimulus by choosing different functions (on the keyboard) that increase/decrease the number of cockroaches; increase/decrease the size of cockroaches; make the cockroaches move; stop the cockroaches; kill the cockroaches (make the cockroaches look dead); and allow the current action to get back to the beginning. The functions can be combined to adapt the exposure exercises to different conditions according to the client's needs. The client observed the AR environment using a Head Mounted Display (HMD) and actively participated in the exposure exercises proposed by the therapist. However, his/her actions (apart from the viewpoint) did not have any direct influence on the AR environment, which was fully controlled by the therapist. The therapist observed the AR environment (client's viewpoint) on a computer screen. For a detailed technical description of the ARET system, see [4].

4.4 Instruments

The intention of the research team was to collect both qualitative and quantitative data. For the quantitative data, different measures were applied to both the clients and the therapists. The clients' clinical outcome measures corresponded to their anxiety, avoidance of the phobic stimulus, and belief in catastrophic thoughts regarding the phobic stimulus (on the 10-point Likert scale). They also performed a Behavioral Avoidance Test (i.e., BAT, adapted from Öst [20]) in order to define (on a 13-point Likert scale) how close the client was able to approach the real cockroach/spider (e.g. 0/12 corresponds to the client refusing to enter the same room where the phobic stimulus is; and 12/12 corresponds to the client interacting with the phobic stimulus for at least 20 seconds). The clients filled out the short version of the Client Working Alliance Inventory [21] questionnaire, which measures (on a 7-point Likert scale) their relationship with the therapist just after the diagnostic interview (before the therapeutic session) and just after the therapeutic session. The therapists filled out a questionnaire, which evaluates the capacity of the ARET system to help them in constructing a therapeutic relationship with the client (as perceived by the therapist). More specifically, the therapists were asked to rate (on a 5-point Likert scale) the degree to which they agreed with statements that described how the system supports (positive valence) or disturbs (negative valence) the therapeutic alliance between the client and the therapist. This questionnaire was based on the short version of the Therapist Working Alliance Inventory [21]. Moreover, the therapists filled out a questionnaire regarding the usefulness and the frequency of use (on a 5-point Likert scale) of the functions available in the ARET system.

For the qualitative data, the therapists responded to an informal interview related to their experience with the ARET system. A video analysis of therapeutic activity was

also applied. In order to preserve the privacy of the participants and not influence the data, all the sessions were recorded using digital cameras and were analyzed (viewed and coded) afterward. No identifiable references to the clients' identities were recorded. More specifically, to record ARET therapeutic activity one video camera recorded a global view and another one focused on the computer screen and the keyboard area. To record IVET therapeutic activity only one video camera was used to record a global view. In addition, commands and interactions of participants with the ARET system were automatically logged during the session. The video analysis covered approximately twenty-five hours of therapeutic session. The following coding scheme for both verbal and non-verbal behaviour was used to capture the distributed activity between the client, the therapist, and their respective interactions (either mediated by technology or non-mediated by technology). The coded activities corresponded to the following categories: verbal communication, visual attention, and performed actions that involved one or both actors and their environment. Once the videotaped sessions were entirely coded, the frequencies of each event were calculated. The reliability of the coding method was assessed by calculating the correlation among 3 different judges that coded in parallel the same 10-minute extract of the video. The correlation was strong ($R=.930$; $p=.002$ between judges 1 and 2; $R=.966$; $p=.000$ between judges 2 and 3; and $R=.944$; $p=.001$ between judges 3 and 1).

Since the therapeutic sessions varied from 1 hour to 3 hours, all data were presented as frequencies per hour. Because some coded activities were not visible during the whole therapeutic session (the camera angle did not allow showing the entire room), the frequencies of the event were estimated according to the calculated frequencies that could be observed for each session. The logs of the commands and interactions of participants with the ARET system could not be used due to technical problems. Therefore, the analysis of the interactions with the system during ARET was performed by video-analysis, and later confirmed by the respective therapists.

4.5 Procedure

All of the clients participated in the diagnostic interview. Once the diagnostic was confirmed, they were informed about the objectives of the study and had to complete a consent form. The clients also filled out the pre-test questionnaires and performed the BAT, after which they received the therapeutic session (following the one-session treatment protocol) using Augmented Reality Exposure Therapy or In Vivo Exposure Therapy. The session was typically organized as follows. The therapists performed the exposure exercises with the clients. These exercises were previously defined during the diagnostic interview and hierarchically organized from the least anxious for the client to the most anxious. Each exposure exercise was first performed by the therapist (i.e., modeling), then the client was invited to repeat the exercise followed by the cognitive restructuring and reinforcement of the therapist. The client-therapist communication regarding anxiety, irrational thoughts, or other issues took place during the therapeutic session. At the end of the session, the clients filled out the same questionnaires and performed the BAT.

4.6 Data analysis

The quantitative data corresponded to the clinical measures and the therapists' answers to the questionnaire. To investigate the clinical effectiveness of each therapeutic session, non-parametrical statistic analysis (Wilcoxon) was applied. In order to explore the potential differences between the traditional IVET and ARET groups regarding clinical outcome measures, the differences between pre and post scores were calculated (delta) for each participant, after which the non-parametrical statistic analysis (Mann-Whitney) was used. In order to study the capacity of the system to help the therapists in constructing a therapeutic relationship with client (as perceived by the therapists), the mean score of four items related to the same dimension was calculated. Afterwards, Cronbach's alpha was applied. Cronbach's alpha was also used to evaluate the agreement of the therapists regarding the usefulness and frequency of use of the functions available in the ARET system. The mean score of four items related to the same dimension was also calculated in order to study the therapeutic relationship between the two actors (as perceived by the clients). Afterwards, Wilcoxon analysis was applied in order to study the evolution of the therapeutic relationship before and after the therapeutic session, and Mean-Whitney analysis was used to compare the therapeutic relationship in two groups. With regard to the qualitative measures, the mean frequencies of events per hour were calculated for each group. The non parametrical statistics (Mann-Whitney) was applied to show significant differences in the therapeutic activity between groups. All analyses were performed using the SPSS 16.0 application with the significance level set at 0.05.

5 Results

5.1 The nature of therapeutic activity

a) Verbal communication

The results in Table 1 show that the verbal communication between the therapist and the client seem to remain similar in both groups ($U=10.00$; $p= .200$ for the therapist verbal communication, and $U=13.00$; $p=.485$ for the client verbal communication). In fact, the therapist-client communication in both groups followed the question-answer model, and the client responded to the therapist's comments.

Table 1. Mean number of verbal communication (VC) per hour and its Mann-Whitney comparison (U) between therapists and clients in each group.

	IVET M (SD)	ARET M (SD)	U (p)
Therapists' VC	86,11 (23,30)	119,32 (42,51)	4,00 (0,14)
Clients' VC	79,54 (24,02)	86,33 (25,33)	7,00 (0,46)

b) Visual attention

The video analysis shows that the client's visual attention seemed to be mainly focused on the phobic stimulus in both groups (virtual phobic stimulus vs real phobic stimulus). However, the clients' visual contact with the therapist significantly differed

between the two groups ($U=.000$; $p=.004$). During the ARET sessions, the clients spent most of the time looking at the virtual animals without having any, or very little, visual contact with the therapist ($M=0,69$; $SD=0,56$ per hour of session). On the other hand, during the IVET sessions, the clients also spent most of the time looking at the real animals, but they had significantly more frequent visual contact with the therapist ($M=37,75$; $SD=10,81$ per hour of session).

The therapists' visual attention was distributed between more sources than the clients' visual attention. In fact, during the therapeutic session, the therapists distributed their visual attention among different targets. The therapists that participated in ARET seem to look mainly at the following targets: (a) visual sources related to the phobic stimulus (i.e., the computer screen, in order to control the client's view of the AR phobic stimulus; the keyboard, in order to choose the appropriate functions in order to expose the client to the phobic stimulus; and the AR marker on which the virtual animals appeared); (b) notes, in order to write comments and/or clinical measures or keep a temporal record of the client's evolution; and (c) the client, in order to see his/her reactions and possible anxiety. The therapists that participated in IVET seem to look mainly at the following targets: (a) the real phobic stimulus, in order to reference themselves to the object in the conversation, or to control; (b) notes, in order to write comments and/or clinical measures or keep a temporal record of the client's evolution; and (c) the client, in order to see his/her reactions and possible anxiety. As Figure 2 shows, the frequency of the therapists' visual attention was significantly higher for the phobic stimulus target and its related visual sources in ARET, and for the notes in ARET, respectively, $U=.000$; $p=.004$; $U=4,000$; $p=.026$; the therapists' visual attention on client comparison was not significant ($U=8,000$; $p=.132$).

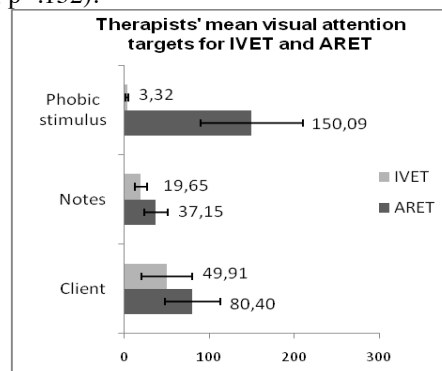


Fig. 2. Mean frequencies of therapists' visual attention per hour in each group.

c) Spatial Orientation

Figure 3 (a) and (b) present the spatial orientation of the two actors. As the results show, the therapist and the client seem to move more in IVET than in ARET. Both actors used all the space available in the therapist's office during IVET, while the actors participating in ARET had limited displacements. With respect to the phobic stimulus, the client and the therapist observed the phobic stimulus in two groups from

two different perspectives (on the table and on the floor). In the ARET sessions, the clients were able to observe the phobic stimulus on the wall as well.

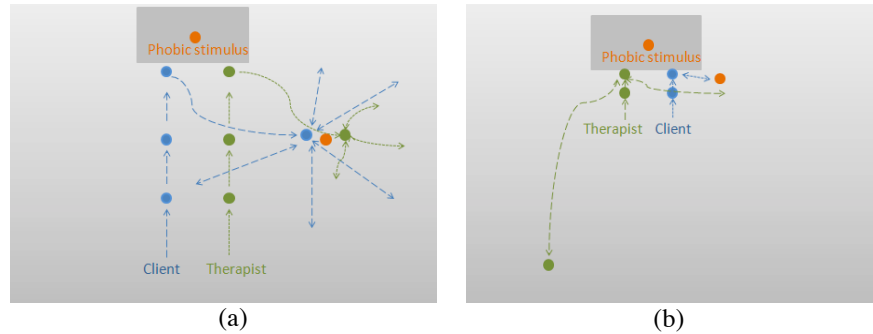


Fig. 3. Example of the spatial displacements of the therapist and client with respect to the phobic stimulus in IVET (Figure 3a) and in ARET (Figure 3b).

d) Interaction with the phobic stimulus

Table 2 shows the variation of exposure exercises observed in both types of therapeutic sessions. The results seem to show that, during ARET, therapists proposed more types of exercises than in the case of IVET. The results also show that, during IVET, the therapists proposed to the clients interact with the real phobic stimulus using a piece of paper or a wooden stick. This was not observed in the ARET sessions.

Table 2. Different exposure exercises proposed by the therapists, observed during the IVET and ARET sessions.

Exposure exercises	ARET	IVET
Observe immobile phobic stimulus	X	X
Observe mobile phobic stimulus	X	X
Observe dead phobic stimulus	X	-
Observe phobic stimulus on personal belongings	X	-
Put hand(s) near phobic stimulus	X	X
Put foot/feet near phobic stimulus	X	-
Interact with phobic stimulus with an artifact (e.g. a stick)	-	X
Kill phobic stimulus	X	X
Throw away phobic stimulus	X	X
Find phobic stimulus under different artifacts	X	-

5.2 Clinical effectiveness

a) Anxiety, avoidance, belief in catastrophic thoughts, and BAT

The analysis of the pre-test shows no significant differences between the IVET group and the ARET group regarding the anxiety measure ($U=13,000$; $p=.415$), the avoidance measure ($U= 12,500$; $p=.373$), the belief in catastrophic thoughts measure ($U=10,000$; $p=.333$), or the BAT measure ($U= 9,500$; $p=.164$). The analysis of the

pre-test and the post-test shows that, for the ARET group, all clinical measures decreased significantly after the therapeutic session ($Z=-2.207$; $p=.027$ for the anxiety measure; $Z=-2.026$; $p=.026$ for the avoidance measure; and $Z=-2.023$; $p=.043$ for the belief in catastrophic thoughts measure). The results also show a significant increase in the BAT scores ($Z=2.232$; $p=.026$) after the treatment (i.e., the clinical improvement on the BAT scale is reflected by an increase of the scores). Also, for the IVET group, the clinical measures decreased significantly after the therapeutic session ($Z=-2.207$; $p=.027$ for the anxiety measure; $Z=-2.014$; $p=.027$ for the avoidance measure). However, for the belief in catastrophic thoughts measure there was no significant difference ($Z=-1.826$; $p=.068$). For the BAT score for the IVET group, the results were significant ($Z=2.232$; $p=.026$).

Figure 6 (a, b, c, and d) shows the comparison analysis of the clinical improvement, which reflects the clinical effectiveness of the two types of therapeutic sessions.

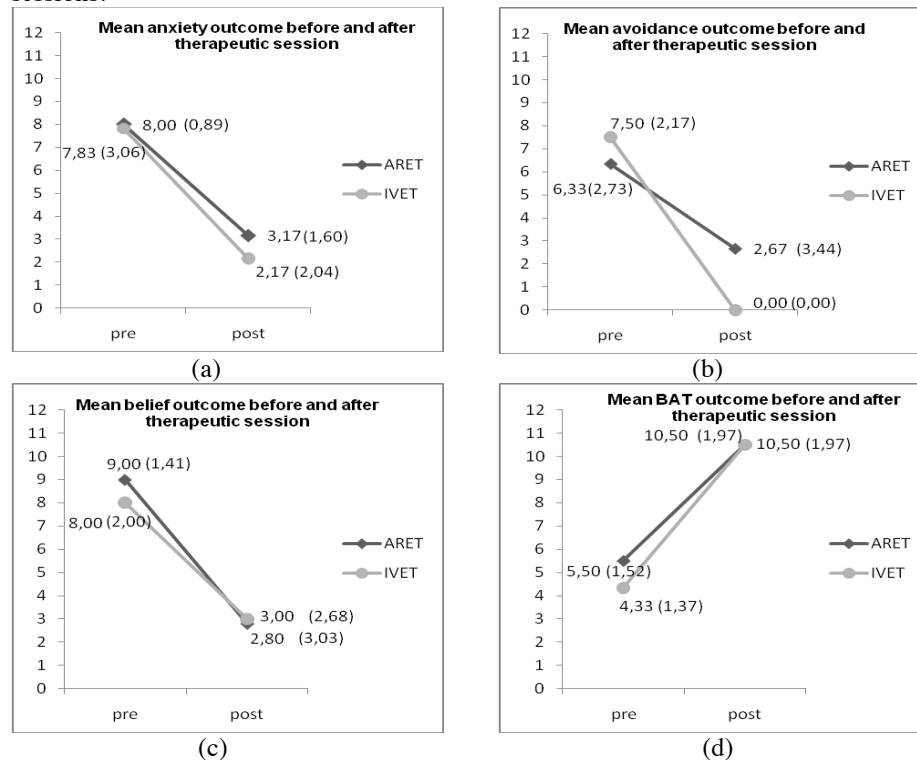


Fig. 6. Mean ratings (SD) for anxiety (Figure 6a), avoidance (Figure 6b), belief (Figure 6c), and BAT (Figure 6d) before and after for both the ARET and IVET sessions.

The results show no statistically significant differences for clinical improvement between the two groups for the anxiety measure ($U= 14,500$; $p=.568$), the belief in catastrophic thoughts measure ($U= 13,000$; $p=.792$), and the BAT measure ($U= 8,500$; $p=.115$). However, the clinical improvement of the avoidance score was significantly higher for IVET ($U= 2,000$; $p=.010$) than for ARET.

b) Therapeutic alliance perceived by the clients

The analysis of the pre-test show no statistically significant differences between the IVET group and the ARET group with respect to the task measure ($U= 12,000$; $p=.328$), the goal measure ($U= 17,500$; $p=.934$), and the bond measure ($U= 15,000$; $p=.625$). This seems to indicate that, in the two groups (during the diagnostic interview), the therapists and clients created similar high therapeutic relationships.

The comparison analysis of the pre-test and the post-test shows no significant differences in the two groups. More specifically, for the ARET group, all three dimensions of the therapeutic relationship were maintained after the therapeutic session ($Z=0.000$; $p=1.000$ for the task dimension; $Z=-0.405$; $p=.686$ for the goal dimension; and $Z=-0.535$; $p=.593$ for the bond dimension). For the IVET group, all three dimensions of the therapeutic relationship were also maintained after the therapeutic session ($Z=-0.736$; $p=.461$ for the task dimension; $Z=-1.095$; $p=.273$ for the goal dimension; and $Z=-1.656$; $p=.098$ for the bond dimension).

Figure 7 (a, b) shows the comparison analysis of the therapeutic relationship before and after the therapeutic session in terms of tasks, goals, and bond. The results show no statistically significant differences between the two groups neither before the therapeutic session (for tasks ($U= 12,000$; $p=.328$), goals ($U= 17,500$; $p=.934$), and bond ($U= 15,000$; $p=.625$)), nor after (for tasks ($U= 12,500$; $p=.373$), goals ($U= 8,000$; $p=.094$), and bond ($U= 12,000$; $p=.315$)).

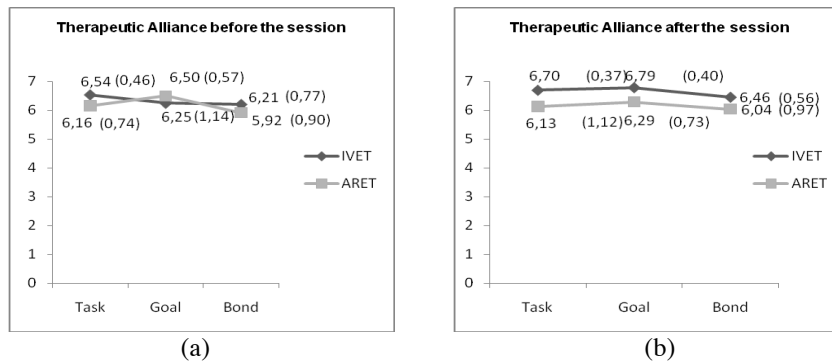


Fig. 7. Mean ratings (SD) for therapeutic relationship in terms of task and goal definitions and bond between the therapist and the client, before (Figure 7a) and after (Figure 7b) the therapeutic session for both the IVET and ARET sessions.

5.3 Acceptability

a) Frequency of use and usefulness of the ARET system

The results show that the therapists seemed to be in strong agreement ($\alpha=.857$) about using all the functions of the ARET system very frequently ($M=4,88$ out of 5; $SD=0,17$) and seemed to be in less agreement ($\alpha=.573$) about these functions being very useful ($M=4,75$ out of 5; $SD=0,39$). The therapists seemed to be in strong agreement that the function “increase the number of cockroaches by 20” was neither

frequently used ($M=1,67$ out of 5; $SD=0,58$) nor useful ($M=2,00$ out of 5; $SD=0,00$). The low score obtained by this function was explained in the informal interview by the fact that the therapists preferred to use an exposure exercise that was closer to real life (a small number of cockroaches or spiders).

b) The appeal of the ARET system to therapist

The results regarding the appeal questionnaire based on the short version of the Therapist Working Alliance Inventory [21] show how, in the therapists' opinion, the ARET system helps them to create the therapeutic relationship with the client in terms of therapeutic alliance (i.e., task, goal and bond). The general internal consistency of the scores given by the therapists was high ($\alpha=.961$). Therefore, the results seem to show that the therapists agree that the system helps the clients in performing therapeutic tasks ($M=3.92$ out of 5; $SD=1.23$), and strongly helps in defining the goals of the therapeutic session ($M=4.42$ out of 5; $SD=0.80$). The results also seem to show that, in the therapists' opinion, the system moderately helps in creating and maintaining the bond between the client and the therapist ($M=3.42$ out of 5; $SD=1.46$).

c) Usability issues of the ARET system

The informal interviews with the therapists as well as the video analysis brought to light some interesting information about the ARET system. According to the therapists, the ARET system itself and its functions were useful for this type of therapy. More specifically, the therapists were satisfied with the system, particularly regarding the sense of controlled and secure context that the ARET system gave their clients. However, the video analysis showed that the therapists used different strategies while performing the same exposure exercise (using the same functions). After deeper analysis of this issue with therapists, we came to the conclusion that the user interface dialogue of the ARET system was not optimal. More specifically, the functions used by the therapists to present the exposure exercises to the client were not optimal. The first usability issue corresponded to the function "Initiate", which allows the therapist to go back to the beginning of the exercise. This function was also identified as a necessary step before increasing the number of cockroaches/spiders. This had its consequences in terms of cockroach/spider movement. Since the "Initiate" function stops the movement, the therapist had to press the "Move" key one more time in order for the phobic stimulus to regain movement.

The second usability issue corresponded to the function that allows the therapist to kill the cockroaches/spiders ("Kill"). This function was used when the clients were smacking cockroaches/spiders with the swatter in order to kill them. The problem appeared when the therapist asked the client to kill more than two animals (i.e., the function worked for only two animals during the same exposure exercise). Once the therapists realized that the "Kill" function only worked for two cockroaches/spiders, they had to find a different strategy to reduce the client's frustration and the possible clinical consequences of not being able to kill all of their phobic stimuli (i.e., using the "stop" functions or "initiate" function).

The third usability issue corresponded to the visibility of the dark animal on the black AR marker. Two techniques were identified as a solution for this concern. The first technique was to make the animal larger by using the "Increase the size" function

in order to see the cockroach/spider outside of the black part of the AR marker. The second technique consisted of moving a cockroach/spider a few centimeters by pressing the “Move” function and then stopping it outside of the AR marker by pressing the “Stop” function. The therapists found these usability issues confusing, making the ARET system less intuitive. They also considered them to be time-consuming.

The therapists recommended several improvements for the system, such as more unpredictable trajectories of animals, additional behavior (e.g. flying cockroaches), incorporation of sounds (e.g. during walking, killing), incorporation of different varieties of animals (there are currently only 3 types of spiders, and 1 type of cockroach). Finally, the therapists mentioned the need for lighter and smaller HMD to reduce the clients’ level of tiredness.

6 Discussion and conclusions

This study evaluates a technology-mediated therapeutic activity (Augmented Reality Exposure Therapy) and a non-mediated by technology therapeutic activity (In Vivo Exposure Therapy) in real-world conditions. The main findings and their implications are discussed below.

The first objectives of this paper were to study how the ARET system supports the therapeutic relationship between the client and the therapist and to determine if the presence of the technology influences the clinical outcome. By applying the multidisciplinary approach, we aimed to demonstrate that, in both technology-mediated therapeutic sessions (ARET) and non-mediated by technology therapeutic sessions (IVET), the relationship between the client and the therapist were the same. More specifically, the results seemed to show that the therapeutic alliance score (evaluated by clients) was similar in both groups after the diagnostic interview (i.e., pre-test measure) and remained similar after the therapeutic session (i.e., post-test measure). Therefore, even though some hesitant clinicians believe that the therapeutic alliance is at risk due to the introduction of technology to the therapeutic process [22], this pilot data is a first step in demonstrating that the therapeutic alliance can be created very well in both treatment conditions and that there are no significant differences between them. The clinical outcome seems to confirm these results. Indeed, the clinical symptoms (i.e., anxiety and avoidance) significantly decreased after the therapeutic session, and the BAT score significantly increased in both groups. Moreover, there was no significant difference between the two groups. The only differences were observed in the avoidance score, which decreased significantly more in the IVET group than in the ARET group. This result might be due to the small sample size of the clinical population used in this study. This assumption can be explained by the fact that these results can be contrasted with direct confrontation to the real phobic stimulus (BAT score). The BAT score results demonstrate that, in both groups, the clients did not avoid the real phobic stimulus after the therapeutic session, but rather interacted with it. In our opinion, the lack of significant differences for the belief in catastrophic thoughts measure in the IVET group is also related to the small size of the clinical sample.

Although some usability issues of the ARET system were detected (i.e., non-optimal user-interface dialogue), overall the therapists evaluated the system as being particularly useful and acceptable. More specifically, the results seem to show that the ARET system provides strong insight into the client-therapist relationship in terms of task and goal definition. In our opinion, the main reason for this conclusion is the flexibility of the system, which proposes various exposure exercises to be presented from different perspectives (on the floor, on the table, and on the wall) to each client, and which allows tasks and goals to be followed. We hope that with the introduction of the changes required by the therapists and improvements in the user-interface dialogue, the system will also provide strong insight into the client-therapist relationship in terms of bond.

The second objective of this paper was to study how the technology-mediated therapeutic activity differs from non-mediated by technology therapeutic activity. The study shows the importance of the distributed, mixed perspective approach in the analysis of this specific collaborative work activity. Indeed, addressing these issues requires observing the activity of the actors in the context of real therapeutic situations. The usual HCI methods that are used to evaluate VR and/or AR applications (e.g. inspections, verification of design guidelines, testbed evaluations) are limited in MH applications in the following way. First, the MH technology evaluation should include day-to-day therapeutic practice. This involves the participation of the clinical population, which is fragile and heterogeneous. Second, the objective of the evaluation should not only consider performances related to the specific sub-tasks (i.e., clinical efficacy), but a larger context related to the client-therapist relationship. Finally, the evaluation involving the clinical population has some clinical and/or ethical constraints that should also be taken into account.

The results show several interesting differences between the ARET and IVET sessions that can be interpreted in terms of future MH technology design guidelines. First, the verbal communications between the clients and the therapists seemed to be the basis of the therapeutic session in both groups. This is understandable since both cognitive reconstruction and reinforcement are one of the most important parts of the phobia treatment. Second, the visual attention of the therapists seemed to differ from the clients' visual attention in the same way in both groups. The therapists frequently changed their visual source of information while the clients focused mainly on one visual source (phobic stimulus). This difference is understandable since the role of both actors in the therapeutic process is different. However, we noticed that the therapists from ARET had more visual attention targets regarding the phobic stimulus (i.e., computer screen, keyboard, and AR marker on which the VR animal appears), and spent significantly more visual attention on it than the therapists from IVET (i.e., real phobic stimulus). Since frequent and numerous switches between different tasks makes the process more demanding for the user and might be a sign of over-load [23], reducing the sources in ARET should be considered in order to make more efficient use of the therapists' resources. Moreover, the clients from the ARET group seemed to have significantly lower visual contact with the therapist than in the IVET group. This result might be explained by the fact that the clients in the ARET group wore HMD, which limited their visual field and head movements. However, even though visual contact plays an important role in all types of face-to-face communication, and its use demonstrates engagement as well as attention and liking (e.g. [24]). The

limited amount of visual contact does not seem to influence the clinical effectiveness of the ARET or the client-therapist relationship. The results also seem to show that ARET differs from IVET in terms of spatial displacement. Indeed, the limited spatial movements in the ARET group can be due to the numerous cables and movement limitations that the system places on the user. However, the displacements in IVET were not only due to fewer constraints on the client, but also due to the therapist trying to control the live phobic stimulus and the client getting more anxious. This might explain why the displacements were so numerous. Finally, the activity analysis seems to show that the ARET system allows therapists to propose a wide range of varied exposure exercises, which in turn allows the client to be exposed to numerous real-life situations that would be difficult to propose with a real, less controllable phobic stimulus. This is an important advantage that can still be improved by the introduction of more direct interactions with the virtual phobic stimulus, such as those present in IVET sessions (i.e., direct interaction with a wooden stick or a piece of paper). Direct interactions of this type can also be developed in terms of simple exposure therapy games, during which the client can interact with the phobic stimulus in a more fun context. In fact, in their recent pilot clinical study, Botella et al. [25] showed the positive effect of a mobile therapeutic game on the client's clinical measures in a case study of cockroach phobia treatment. Thus, the enjoyment factor related to the exposure seems beneficial for clients and can open up a new range of related research.

The results show that applying both a multidisciplinary and mixed perspective approach can bring to light a great amount of interesting information for future design. In fact, this clinical setting evaluation has identified several shortcomings that highlight the need to improve both the design and the evaluation methods of the MH technologies. First, there is a need to maximize non-clinical, usability evaluation. Indeed, even though the ARET system involved a collaborative design process in which both MH professionals and designers were involved, some usability issues were only detected once the finished ARET system was tested in a clinical setting. Thus, an evaluation using scenario-based tests to detect all possible interaction issues should be performed to help improve requirements gathering. Second, there is a need to take into account all the actors participating in the therapeutic process. While the therapist plays a very important role in the design and evaluation process [1], the therapeutic activity should be analyzed by taking into account both the client and the therapist and their respective and related interactions. More specifically, the interactions between these two actors should first be understood in the traditional clinical setting, and then translated to the technology-mediated setting in order to support these interactions. Similarly, in our opinion, the comparison evaluation such as the one presented in this study is beneficial for future design recommendations. Third, the introduction of more natural user interfaces and the development of innovative interaction metaphors that allow more natural and direct client-therapist interactions (i.e., without HMD, cables, computers screen) should be applied. According to Riva et al. [27] in order to allow the technology-mediated therapeutic process to be effective, the mediation of the technology should disappear from the client's awareness (*disappearance of mediation*). When this phenomenon occurs, the user is not simply observing the projected virtual environment but is actively participating in the therapy. The therapist-client interactions are one of the main factors in creating this phenomenon. Also, therapist-client interactions create *common*

ground between the two actors, which is so important to clinical experience [26]. Finally, the design of the system should take into account a large clinical population. Even though the ARET system is presented in a case study for cockroach and spider phobia, its advantage is based on the fact that the simple introduction of different virtual objects such as snakes, bats, or rats would allow the system to treat different small animal phobias. This flexibility should be considered in all MH applications.

This study can be improved in several ways. First, all the conclusions must be confirmed with a larger clinical population sample. Second, in addition to the frequency of events, additional measures such as time would be interesting to consider. Finally, the appeal of the ARET system to the clients in term of therapeutic alliance could also be evaluated. Even though this preliminary study has limitations, the evaluation process gave us a lot of interesting information in terms of design and evaluation guidelines and showed the importance of taking into account the multidisciplinary and mixed perspective approach in new MH technologies evaluation.

Acknowledgements

We would like to thank to all therapists and clients that participates in this study. We would also like to express a special gratitude to Juana Maria Breton Lopez and Patricia Mesa Gresa for their valuable help. This study was funded by Ministerio de Educación y Ciencia Spain, Project Game Teen (TIN2010-20187) and partially by projects Consolider-C (SEJ2006-14301/PSIC), “CIBER of Physiopathology of Obesity and Nutrition, an initiative of ISCIII” and Excellence Research Program PROMETEO (Generalitat Valenciana. Conselleria de Educación, 2008-157).

References

1. Doherty, G., Coyle, D., Matthews, M.: Design and evaluation guidelines for mental health technologies. *Interacting with Computers*. 22(4), 243--252 (2010)
2. Coyle, D., Doherty, G., Sharry, J., Matthews, M.: Computers in Talk-Based Mental Health Interventions. *Interacting with Computers*. 19(4), 429--586 (2007)
3. Coyle, D., Doherty, G.: Clinical evaluation and collaborative design: developing new technologies for mental healthcare interventions. *Proc. CHI*. 2051--2060 (2009)
4. Botella, C., Juan, M., et al.: Mixing Realities? An application of Augmented Reality for the treatment of cockroach phobia. *CyberPsychology & Behavior*. 8(2), 161--171 (2005)
5. Botella, C., Bretón-López, J.M., Quero, S., Baños R.M., García-Palacios, A.: Treating Cockroach Phobia With Augmented Reality. *Behavior Therapy*. 41(3), 401--413 (2010)
6. Brinkman, W.P., Sandino, G., van der Mast, C.: Filed observations of therapists conducting virtual reality exposure treatment for the fear of flying. *Proc. ECCE* (2009)
7. Paping, C., Brinkman, W.P., van der Mast, C.: An Explorative Study into a Tele-delivered Multi-patient Virtual Reality Exposure Therapy System. *Wounds of War II*, IOS press, Amsterdam, The Netherlands. 203 - 219 (2010)

8. American Psychiatric Association: *Diagnostic and Statistical Manual of Mental Disorders*, Fourth edition, Text revision. Washington, D.C.: American Psychiatric Association. (2000)
9. Alonso J, Angermeyer M.C., Bernert S., et al.: Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatrica Scandinavica*. 109(420), 21--27 (2004)
10. Van Hout W.J.P.J., and Emmelkamp P.M.G.: Exposure in Vivo Therapy. *Encyclopedia of Psychotherapy*. 761--768 (2003)
11. Choy Y, Fyer AJ, Lipsitz JD.: Treatment of specific phobia in adults. A comprehensive review on the treatment of specific phobia. *Clin Psychol Rev* 27, 266--286 (2007)
12. Powers, M.B., Emmelkamp, P.M.G.: Virtual reality exposure therapy for anxiety disorders: A meta-analysis. *Journal of Anxiety Disorders*. 22(3), 561--569 (2008)
13. Baños, R.M., Botella, C., Garcia-Palacios, A. et al.: Presence and reality judgment in virtual environments: a unitary construct? *CyberPsychology & Behavior*. 3(3), 327--335 (2000)
14. Horvath, A.O.: The Alliance. *Psychotherapy*. 38(4), 365--372 (2001)
15. Assay, T.P., Lambert, M.J.: The empirical case for common factors in therapy: Quantitative Founding. In: Duncan, B.L., Hubble, M.L. & Miller, S.D. (Eds.). *The Heart and Soul of Change*, Washington, DC, American Psychology Association. 23--55 (1999)
16. Meyerbröker, K. and Emmelkamp, P.M.G.: Therapeutic Process in Virtual Reality Exposure Therapy: The Role of Cognitions and The Therapeutic Alliance. *Journal of Cybertherapy & Rehabilitation*. 1(3), 247--257 (2008)
17. Nardi, B., Kaptelinin, V.: *Acting with Technology: Activity Theory and Interaction Design*. MIT Press. (2006)
18. Hollan, J.D., Hutchins, E.L.: Opportunities and Challenges for Augmented Environments: A Distributed Cognition Perspective. In: *User Friendly Environments: From Meeting Rooms to Digital Collaborative Spaces*. Lahlou, S., Springer Editor. (2009)
19. Wrzesien, M., Burkhardt, J.M., et al.: Analysis of Distributed-Collaborative Activity during AR Exposure Therapy for Cockroach Phobia. *Proc. Cyberpsychology* (2010)
20. Öst, L.G.: Rapid treatment of specific phobias, In: Davey, G.C.L. (Eds.). *Phobias: A handbook of theory, research, and treatment*. New York: Wiley. 227--247 (2000)
21. Tracey, T.J., Kokotovic, A.M.: Factor structure of the Working Alliance Inventory. *Psychological Assessment*, 1, 207--210 (1989)
22. Germain, V., Marchand, A., Bouchard, S., Guay, S., Drouin, M.S.: Assessment of the Therapeutic Alliance in Face-to-face or Videoconference Treatment for Posttraumatic Stress Disorder. *Cyberpsychology, Behavior, and Social Networking*. 13(1), 29--35 (2010)
23. Neerinx, M.A., van Besouw, N.J.P.: Cognitive task load: a function of time occupied, level of information processing and task-set switches. *Engineering Psychology and Cognitive Ergonomics*. 31 (6), 247--254 (2001)
24. Piper, A.M., Hollan, J.: Analyzing Multimodal Communication around a Shared Tabletop Display. *Proceedings of ECSCW*. 283--302 (2009)
25. Botella, C., Breton-López, J., Quero, S., Baños, R.M., Garcia-Palacios, A. et al.: Treating cockroach phobia using a serious game on a mobile phone and augmented reality exposure: A single case study. *Computers in Human Behavior*. 27(1), 217--227 (2011)
26. Riva, G., Zurloni, V., Anolli, L.: Client-Therapist Communication in Computer assisted environment. In: Anolli, L. et al.: *The Hidden Structure of Interaction: From Neurons to Culture Patterns*, Amsterdam: IOS Press, (2005)