1 2 9 0

UNIVERSIDADE D
COIMBRA

Marta Sofia Paiva Nunes

Exploring Augmented Reality for Exposure Therapies

Dissertation submitted to the Department of Electrical and Computer Engineering of the Faculty of Sciences and Technology of the University of Coimbra in partial fulfillment of the requirements for the Degree of Master of Science oriented by Professor Doctor Paulo Jorge Carvalho Menezes

January of 2023

1 2 9 0

UNIVERSIDADE D
COIMBRA

Marta Sofia Paiva Nunes

Exploring Augmented Reality for Exposure Therapies

Thesis submitted to the
University of Coimbra for the degree of
Master in Electrical and Computer Engineering

Supervisor:
Paulo Jorge Carvalho Menezes

Coimbra, 2023

This work was developped in collaboration with:

University of Coimbra

Department of Electrical and Computer Engineering

Institute of Systems and Robotics

Dedication

Apesar de não ser muito extensa, a realização desta dissertação é o culminar de muito esforço, stress e dedicação e o fim de uma jornada louca mas gratificante. Uma jornada destas, que demorou tanto tempo e trouxe tantos desafios, não poderia nunca ter sido feita sozinha. A todos os que me apoiaram, acompanharam e estiveram comigo durante este tempo, esta tese é dedicada a vocês. Ao meu professor e orientador Paulo Menezes, por me guiar neste caminho e ajudar-me a alcançar os resultados finais que pretendia. Aos meus colegas do IS3L, pela ajuda nos problemas que surgiam, pelas conversas de motivação ou de aconselhamento e pelos momentos de distração. A vossa ajuda e companhia foi essencial para conseguir manter a motivação, principalmente em alturas mais complicadas. Aos meus colegas do mestrado, Diogo, Marco, Óscar e Miguel, por todo o companheirismo e ajuda nos momentos mais difíceis. Aos colegas que estão comigo desde o início do curso Inês, Mara, Paiva, Luís e ao Matias, que me acompanha desde antes disso, por todos os dias e as noites, por toda a animação e toda a amizade. A todos os que durante esta jornada pude considerar amigos, por terem animado e melhorado esta jornada. Ao Gonçalo, pela paciência infinita, por estar sempre presente e por todo o amor. À Raquel e à Bia, por saber que posso sempre contar convosco e por me conhecer quase melhor que ninguém. A todos os meus tios e primos, por todos os sábios conselhos, dicas e palavras de incentivo. Aos meus irmãos por toda a parvoíce infinita que parecem conter. Por fim, aos meus pais, por sempre me apoiarem e suportarem em todas as decisões desta jornada, guiando sempre e ensinando-me o mais importante. Sem eles, nada disto era possível e espero que estejam tão orgulhosos como eu pelo fim desta etapa.

Marta Nunes

Acknowledgments

The author wants to sincerely thank Drª Alexandra Pais for her insight into the work done and the discussion of the uses of technology in therapy.

These special thanks are also extended to all the participants of the tests made, all of whom were very helpful, and attentive during its conduct.

The author thanks the Institute of Systems and Robotics (ISR) for providing the resources necessary during the development of this work.

Abstract

In vivo exposure is the most common treatment for phobias. However, it has several drawbacks that can be mitigated by adopting technological alternatives, such as virtual or augmented reality. Augmented reality provides some advantages over virtual reality, including fewer modelling costs and a higher level of realism. As a result, this dissertation aims to develop an alternate treatment to exposure *in vivo* using augmented reality with Hololens 2. This solution is divided into two parts: the therapist's computer and the augmented reality device used by the patient. On the therapist's side, they can add, move, and remove phobic elements. It is also possible to choose the type of element to add, its size, and its starting position. It is also possible to observe the therapy from various points of view and the patient's interaction with the virtual objects. On the patient side, the phobic elements added by the psychologist are visible. The proposed approach allows the patient to interact with the phobic elements while simultaneously giving the psychologist complete control over them, allowing each person to have a unique and personalized experience based on their phobias. The psychologist then gets a tool to aid in the therapy that allows him to decrease the time and difficulty of preparing the therapeutic session, while also giving increased confidence, since the phobic elements do precisely what is asked of them. These features allow for a greater focus of the therapy on the patient and their reactions to the presented elements. To assess the implementation, three tests were performed: one to evaluate the usability of the application and two for the realism of the models used. The final result of the work was also discussed with a therapist to study its usability in the therapeutic context. The results were overall positive, proving the usefulness of the systems in the aid of treatment for phobias.

Resumo

A exposição *in vivo* é o tratamento mais comum para as fobias. Contudo, tem vários inconvenientes que podem ser mitigados pela adopção de alternativas tecnológicas, tais como a realidade virtual ou aumentada. A realidade aumentada oferece algumas vantagens sobre a realidade virtual, incluindo menos custos de modelação e um nível mais elevado de realismo. Como resultado, esta dissertação visa desenvolver um tratamento alternativo à exposição *in vivo* utilizando a realidade aumentada com Hololens 2. Esta solução divide-se em duas partes: o computador do terapeuta e o dispositivo de realidade aumentada utilizado pelo paciente. No lado do terapeuta, este adiciona, move e elimina os elementos fóbicos. É ainda possível escolher o tipo de elemento a adicionar, o seu tamanho e a sua posição inicial. É ainda possível observar a terapia de vários pontos de vista e a interação do paciente com os objetos virtuais. No lado do paciente, são visíveis os elementos fóbicos adicionados pelo psicólogo. A abordagem proposta permite ao paciente interagir com os elementos fóbicos ao mesmo tempo que dá ao psicólogo um controlo completo sobre eles, permitindo que cada pessoa tenha uma experiência única e personalizada baseada nas suas fobias. O psicólogo obtém então uma ferramenta de auxílio à sua terapia que lhe permite diminuir o tempo e a dificuldade de preparação da sessão terapêutica, dando ainda uma confiança acrescida, uma vez que os elementos fóbicos fazem exatamente o que lhes é pedido. Estas funcionalidades permitem um maior foco da terapia no paciente e nas suas reações aos elementos apresentados. Para avaliar a implementação, foram realizados três testes: um para avaliar a usabilidade da aplicação e dois para o realismo dos modelos utilizados. O resultado final do trabalho foi também discutido com um terapeuta para estudar a sua usabilidade no contexto terapêutico. Os resultados foram globalmente positivos, provando a utilidade dos sistemas na ajuda ao tratamento de fobias.

List of Figures

List of Tables

Acronyms

AR augmented reality.

ARET augmented reality exposure therapy.

HMD Head Mounted Display.

ID identification.

IP Internet Protocol.

IR Infrared.

LAN Local Area Network.

MRTK Mixed Reality Toolkit.

PC personal computer.

QR Quick Response.

RPC Remote Procedure Call.

TCP Transmission Control Protocol.

TL Therapeutic Lamp.

ToF Time-of-Flight.

UDP User Datagram Protocol.

UEQ User Experience Questionnaire.

USB Universal Serial Bus.

VR virtual reality.

VRET virtual reality exposure therapy.

Contents

1

Introduction

Anxiety affects 18.1% of adults, making it one of the most common mental diseases [1]. When at excessive levels, it causes agitation, muscle tension, and difficulty concentrating, among other symptoms. Some anxiety disorders are associated with phobias, which are aversions or fears of specific items or circumstances that cause significant discomfort to the individual who comes into contact with them [2][3].

There are various treatments for these diseases, and while no single treatment works for all phobias, *in vivo* exposure is the most robust treatment applicable to most of them [4]. In this type of treatment, the patient directly confronts his phobic stimulus to overcome it. This therapy is successful because there is a desensitization to the phobic element, that is, the patient gets used to the phobic element, which allows him/her to tolerate it on a daily basis. However, this therapy has many problems associated with it, such as low acceptance and high dropout rates [14, 16], complications related to the presence of the phobic element (especially if it is an animal) [11, 12, 14], the impossibility of performing this therapy in the usual office environment [7], and the therapist's inability to control the phobic elements.

Technological solutions have been introduced in therapies using virtual and augmented reality to overcome the drawbacks of *in vivo* exposure. As virtual reality (VR) creates simulations of the real-life feared environment/elements, several studies have found that using it can be just as successful as using *in vivo* exposure [4, 5, 6, 7, 8, 13, 14]. On the other hand, being relatively new in this application field, through augmented reality (AR) the patient may see the real world with phobic elements added. Despite being recent, this treatment method yields positive results compared to *in vivo* exposure [6, 9, 10, 11, 12].

Both treatments minimize the problems discussed above and may allow therapists to have complete control of the phobic elements, allowing gradual and tailored exposures based on patients' anxieties [5, 9, 10, 11, 14]. Plus, technological solutions offer a wider variety of exercises and scenarios to the patient [11, 12].

In the case of exposure therapy applications, AR has some advantages over VR, which goes from less demanding modelling needs, to improved user experiences. As a matter of fact, in augmented reality, the environment does not need to be created, and as such, it has lower programming and modelling requirements [6, 11, 12]. On the other side, the integration of virtual items in the actual world contributes to a more natural integration and acceptance from the user's point of view, also contributing to a stronger sense of presence [9, 11].

However, both virtual and augmented reality demand caution in their application, as they aim to produce experiences that convey sensations comparable to those in the real world. If successful, they produce realistic impressions which may be, in turn, susceptible to inducing emotional involvement, discomfort levels, and even anxiety [10].

We can better understand which applications are best for VR or AR by simply considering that the first "transports the user to another virtual place". In contrast, the second one brings virtual elements to the user space, or eventually modifies some of the existing ones. From this, it is simple to understand that virtual reality is particularly appropriate for cases such as claustrophobia and acrophobia. On the other hand, augmented reality can be very effective for animal phobias and other elements that can be integrated naturally into the user's surroundings and familiar environment, without the risk of cybersickness.

1.1 Related Work

The first alternatives to *in vivo* exposure have emerged using virtual reality. The work of Rothbaum *et al.* [20] (1995) was one of the first to present an experiment with VR using a headset. It proposed a treatment for acrophobia using various virtual environments: various bridges, outdoor balconies at different heights, and one elevator simulation. Although there was no treatment comparison group and the subjects were not patients seeking treatment, the study successfully showed that virtual reality could be used to treat anxiety disorders.

In Côté and Bouchard's work [21] (2005), the authors focus on treating arachnophobia with a VR environment consisting of two apartments with many rooms with different levels. There were a few static spiders in the lower levels, while the ones in the most difficult levels were bigger and moved. The study showed improvements in participants' phobia symptoms.

Pitti *et al.* [22] (2015) presented an alternative treatment to agoraphobia with VR, using polarized glasses and projecting the image on a special screen. In this experience, there were various scenarios: a square and street, an airport building and plane, a bank office, an elevator and underground car park, a beach, a highway, and a cableway. This study concluded that the patients which received combined treatment (VR and psycho drugs) improved significantly more than the others. However, the dropout rates with VR and *in vivo* exposure were similar.

The work of Toma *et al.* [29] (2020) intends to treat the phobia of snakes through a 3D game with two scenarios. In this game, the player has to collect coins while avoiding or killing snakes that crawl and try to bite, given that the two scenarios have different levels of difficulty. This solution was able to decrease the fear of snakes among the participants. However, the authors proposed an improvement that enables the game to give an adaptable degree of intensity of exposure.

As explained before, as well as virtual reality, augmented reality solutions are also used in a therapeutic context.

Juan *et al.* [23] (2005) presented one of the first attempts to use AR as an alternative to *in vivo* exposure, using a VR Head Mounted Display (HMD) and a camera connected to a computer. This work, although exploratory, already supported the idea of having the therapist watch the exposure scenario on the computer monitor. The virtual elements used were spiders (three types: small, medium, and large) and cockroaches (one type), and these could move, be static, or even dead, if the patient killed them. The interaction provided to the therapist was based on markers, controlling, thus, where the objects would appear, how many animals appear, their size, and movements.

Bretón-López *et al.* [25] (2010) developed an experience using an AR system to treat cockroach phobia. The system ran on a personal computer (PC), and similarly, the video was captured by a camera attached to an HMD and connected to the PC via Universal Serial Bus (USB). The therapist could control the cockroaches' number, movement, and size. The experience reported good results, however, the intricate cabled connection caused discomfort and tiredness.

The work of Wrzesien *et al.* [11] (2011) and Botella *et al.* [12] (2016) also aimed to treat spider and cockroach phobias using AR solutions. The patients could observe and interact with the scene through HMDs, and the therapists controlled the phobic triggers' sizes, number, and movement. Despite the success of the therapy [11], it is mentioned that the systems have functions (like the non-optimal user-interface dialogue) that are not useful or not optimal, and [12] tiredness and dizziness were reported due to the HMD.

Fatharany *et al.* [31] (2016) proposed an application that replaces the markers usually necessary in AR systems with everyday objects. This uses a camera placed above the desired scene, whereby the patient sees the scene on a computer. The cockroaches used in the application interact with the objects and appear from underneath. However, the patient has to look to the monitor to acknowledge the animals, which interferes with the sense of presence desired.

Plasencia and Escobar-Sánchez [30] (2018) developed a system to mitigate arachnophobia through augmented reality. This system uses a mobile device or a computer and places spiders on markers, positioned on or close to the patient. This application was able to decrease the fear of spiders, however, the design and animation of the animals could be improved and the use of markers affected the experience.

The work of Tripi *et al.* [27] (2018) presented a virtual reality exposure therapy (VRET) and augmented reality exposure therapy (ARET) framework for phobia treatment. This framework was designed to be easily customized, allowing the creation of different scenarios. The system consists of the therapist's computer where the environment was configured and controlled, and a mobile application that allowed the patient to see the virtual environment (in the case of VR), or see the real world with the phobic elements added (in the case of AR).

De Witte *et al.* [10] (2020) focused on ensuring the validity of using an AR application for smartphone or tablet to induce fear and anxiety, an important prerequisite for effective treatments with AR. The user could choose between various animals (rats, pigeons, snakes, cockroaches, spiders, frogs, and dogs). The study showed that the application induces fear and anxiety, even though the use of a smartphone or a tablet also reduced the realism and perception of the virtual elements, and made it difficult to interact with these elements.

Patrão, Menezes and Gonçalves [24] (2020) presented an AR experiment with spiders. The application used tablets, one for the therapist and another for the patient, and both could control and add elements to the environment. The study revealed good usability and interaction, yet, the use of tablets reduced the perception of the virtual elements and made it difficult the interaction with them.

The work of Corbett-Davies, Dünser and Clark [28] (2012) described an AR system for exposure therapy, using Microsoft Kinect to see the scene and a monitor to show it with the phobic elements added to the patient. Despite the animals (spiders, in this case) being realistic, the fact the AR scene was viewed on a computer screen affected the immersion of the user, that could look away from the monitor to stop seeing the scene.

Wrzesien *et al.* [26] (2013) approached AR therapy with a new device, the Therapeutic Lamp (TL), which projected the elements on a surface. This experiment focused on cockroach and spider phobias, with three types of each animal, and the therapist controlled the number, size, and movement of the elements. The experience showed that the TL is an effective tool for small animal phobia treatment, however, the fact that it used projection instead of holograms could affect the sense of presence and realism.

These studies in the field of virtual reality and augmented reality have shown the potential for their use in therapeutic contexts. While VR has been used in therapy for a while, AR has emerged as a promising solution in cases where it is advantageous to add virtual objects to the real world, such as in the treatment of animal phobias.

The increasing interest in AR therapy is also due to the availability of usable AR headsets, which allows for a more familiar and less anxiety-inducing environment compared to VR environments. However, after reviewing the literature, it was found that the devices used in many AR therapy approach, such as tablets, smartphones, and TL, were not always the most appropriate. These devices can impact the realism of therapy elements and the sense of presence, and previously used HMDs can be uncomfortable and cumbersome due to their weight and wires.

Our alternative, presented below, aims to provide a more realistic and comfortable augmented reality therapy experience, to further improve the effectiveness of AR in therapeutic contexts.

1.2 Main Objectives

Given the multiple problems inherent to exposure therapy and the characteristics that make it successful, in this thesis, we aim to create an alternative to these problems using AR. For this, the main objective is to develop an application for AR devices that can be used in a therapeutic environment. The therapy then needs two focuses: the patient and the therapist. It is necessary to create a solution that the patient wears comfortably and fits their needs. The therapist needs a simple solution that is not distracting, and allows for the treatment of various phobias. Therefore, we want the solution to meet the following objectives:

- Creation of an AR app to support exposure therapies in the office environment. To be useful, this application must:
  - have an added value for the therapist, being a support tool in their work, since not many exist;
  - provide a realistic experience that helps patients face and overcome their fears;
  - allow the therapist to have complete control of the phobic elements and the levels of exposure to them, allowing the creation of a unique and personalized experience for each person;
  - allow the therapist to analyze the patient and their reactions;

  - to allow the use of a space in which the patient feels comfortable and with someone they trust.

- Validation of the app created in terms of user experience and added value for the therapist.

1.3 Document structure

This dissertation is organized as follows:

- **Chapter 1** presents the motivation for the work and its relevance, the state of the art with VR and AR, and the objectives of the work.
- **Chapter 2** explains the requirements that need to be implemented in order to consider the work successful.
- **Chapter 3** clarifies in detail the proposal and which materials and tools were used in the development of the work and why these were chosen.
- **Chapter 4** describes all the implemented features, presenting the final interface of the work.
- **Chapter 5** delivers all the studies carried out to evaluate the work done and their respective results.
- **Chapter 6** finalizes this dissertation by presenting conclusions about the work done and indicating some further work that would be interesting to be done in order to improve the existing application.

2

Requirements

Any project needs a definition and specification at the beginning of the project phases and requirements that need to be fulfilled in order to be completed successfully. This definition of requirements can be done in various ways, such as through user stories and use cases. To better understand the problem we want to solve and the requirements, let's start by reviewing what kind of therapy is currently performed when no technological alternative is available.

Ana has agoraphobia, which manifests itself through an inability to go to the supermarket. This affects her day to day life, since she always depends on other people to do her weekly shopping. Ana then decided to start therapy. Part of this therapy includes going to the supermarket with the therapist, which involves changing the location of the therapy, choosing a suitable place, and is always conditioned by the number of people present in the store.

John has arachnophobia on a level that causes him to leave a room where he has seen this animal. He has decided to start therapy, and the sessions take place entirely within the clinic. However, before therapy the therapist has to find spiders, feed them, store them, and keep them alive so that they can later be shown to the patient. In addition, the exposure involves the spiders remaining enclosed so that they do not escape or frighten the patient.

Figure 2.1: Examples of therapies without technology.

Figure 2.1 shows two examples of exposure therapy without any technological devices. In the first example, exposure therapy implies that it takes place outside the clinic, requiring travel. These trips require additional preparation by the choice of the ideal location. However, even with all the precautions, unforeseen events may occur during these trips that negatively affect the patient. If we analyze other phobias, such as the fear of flying, a treatment of exposure without recourse to technological means is not possible.

The second example refers to an animal phobia that, depending on the size of the animal in question, may or may not make therapy in the clinical setting possible. Exposure therapy for this type of phobia involves capturing, feeding, and keeping the animal as long as necessary for therapy, which can be extremely difficult depending on the type of animal in question. During therapy, these animals still have to be caged or leashed to prevent them from escaping or approaching the patient, which can have the opposite effect to that intended by the therapy. Another disadvantage related to the involvement of animals in therapy is that they may not only move unexpectedly, but also make noise that affects the patient and worsens his or her condition.

Having understood the problem to be solved, the next step is to define the functional and non-functional requirements, so that we can proceed with development. The functional requirements are the characteristics of a product that the developers have to implement so that the final product fulfills the requirements. The non-functional requirements define how a system should behave, in terms of speed, and usability, among others. In table 2.1 are presented the functional requirements, and the non-functional requirements are presented in table 2.2.

Requirement	Description
1	Choice of the phobic element and its characteristics
1.1	Choose the virtual model to add
1.2	Select the size of the virtual phobic element
1.3	Select whether the animal has animation or not
1.4	Pick the addition position of the virtual model
2	Editing of the phobic element
2.1	Addition of the virtual model of the chosen animal
2.2	Removal of any previously added phobic elements
2.3	Movement of the phobic element to the desired position
3	Possibility of having more than one type of phobic element present in the same session
4	Visualize the patient's interaction with the phobic elements
5	View therapy from the most convenient perspective

Table 2.1: Functional requirements

Requirement	Description
1	Communication between therapist and patient application
2	Synchronism in adding, removing, and moving phobic elements in real time
3	Realism of the phobic elements
4	Simple and clean interface on both the patient and therapist side
5	Usability and user-friendly interface in the therapist application
6	Non-distracting application

Table 2.2: Non-functional requirements

3

Proposal and development materials

Taking into count the necessary therapeutic context for the project to have good results and be well received by both health professionals and patients, we will now present the detailed proposal of the work, the materials, and the development tools used in it.

3.1 Proposal

As we had started discussing before, the goal of this proposal focuses on developing a passive augmented reality solution to be used in a clinical setting to replace *in vivo* exposure. This is expected to enable the therapist to bring into the office space virtual versions of elements associated with specific phobias of each patient. Thus, the main objective is not to replace the therapist or therapy, but to provide a tool to ease the health professionals' work. This system should be integrated into a typical therapeutic office space, where a table can be used in front of the patient, where the exposure elements will appear, and with which the patient will be asked to approach (when deemed appropriate by the therapist), or even interact. This approach may also allow us to explore the surrounding space, floor, walls, and ceiling to introduce these types of elements. For example, it is more natural a dog to appear on the floor than on a table, whereas a spider might be on the wall or on the ceiling.

3.1.1 Putting the therapist in control

One of the innovations introduced by this approach is the possibility that the therapist has to control the whole exhibit through a computer, selecting the elements that should appear, their location and characteristics, among others. Similarly, they can be removed if it is found that there is some overexposure effect. It will also be possible to choose the most appropriate phobic element for each patient, allowing a single device to be versatile for several patients with different phobias. In this way, we avoid the cumbersome work of getting the phobic elements used in the therapy. We also expect to improve the acceptance of this type of therapy of exposition, since there is no danger to the patient of the phobic element being closer or doing something that the individual cannot tolerate.

3.1.2 Integrating the therapist's workspace

One of our objectives is to take advantage of the space where the therapeutic session takes place and integrate the virtual triggering elements in an ecological way. This is to say that the virtual elements must be added in a coherent way to the space and existing furniture. For the therapist to take control, as referred above, there must be a common reference frame between their view and the patient's. This common frame enables the introduction of a virtual table on the therapist's view that corresponds to the physical table viewed by the patient. This allows the clinician to precisely choose where to place the virtual elements on this surface. As the 3D reconstruction of the environment in real-time can require a significant amount of computational power and may result in the appearance of disturbing artifacts, the choice fell on the superposition of models with some of the furniture elements. Thus these elements will be represented in the therapist's view so that they can be used to precisely place those referred elements, while ”invisible” in the patient's, as there is already the physical counterpart. This idea is accomplished with the use of fiducial markers (in this case, Quick Response (QR) codes).

3.1.3 Solution Overview

Given the topics discussed, we now have an idea of the solution we want to create. The application is divided into two parts: the patient's device and the therapist's device. On the patient's side, the user is able to see his phobic triggers through an HMD able to create AR solutions. On the other hand, the psychologist can control everything the patient sees in real-time. This includes managing the type, quantity, size, and movement of the phobic aspects visible to the user at every moment, allowing the creation of experiences tailored to each patient's needs. With the ability to choose the type of phobic trigger presented to the patient, the treatments of numerous phobias will also be possible with the same equipment.

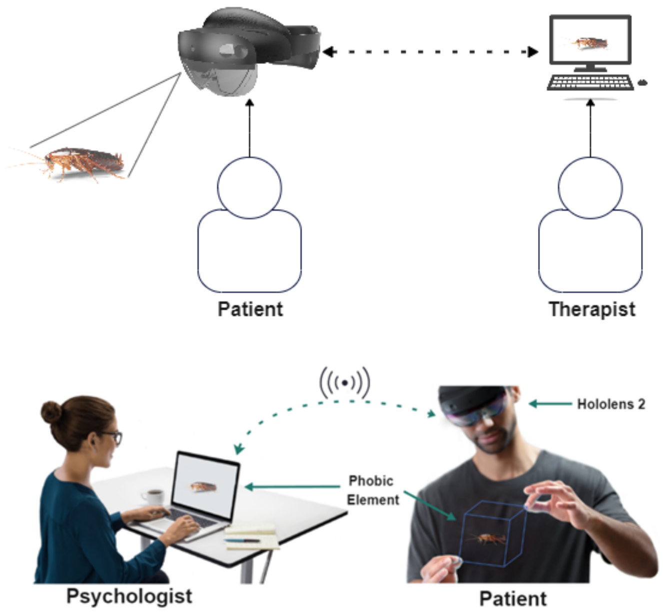

Figure 3.1: Illustration of the principles supported in the application.

Figure 3.1 shows two similar diagrams portraying the application's division into the two parts explained above. The patient has an immersive experience using the HMD, viewing his phobic element in numerous ways. The psychologist controls everything the patient sees via a computer. Being in the same room, the therapist also sees the patient, being able to evaluate his/her reaction to the phobic elements, and adjust them.

Figure 3.2 shows the flow diagram of our solution. The HMD and its sensors can acquire a model of the room, the user's point of view, and perform surface reconstruction. The application includes fiducial marker detection and poses extraction to support the establishment of a common reference frame for both the HMD and desktop applications. The reconstruction of the surfaces gives realism to the scene because the phobic elements can be put realistically on the objects in the room. The user's point of view is constantly monitored and used by the therapist to analyze in real time what is in the patient's field of view and if they try to avoid looking at, or staring at some of the elements. All these features support this therapeutic system composed of the two already mentioned applications. This is made possible through the inclusion of appropriate communication support that will be described later. By putting it all together, the virtual elements may be added, removed, and controlled by the therapist, according to the interaction and intended stimulation for each patient at each specific moment of the exposure sessions.

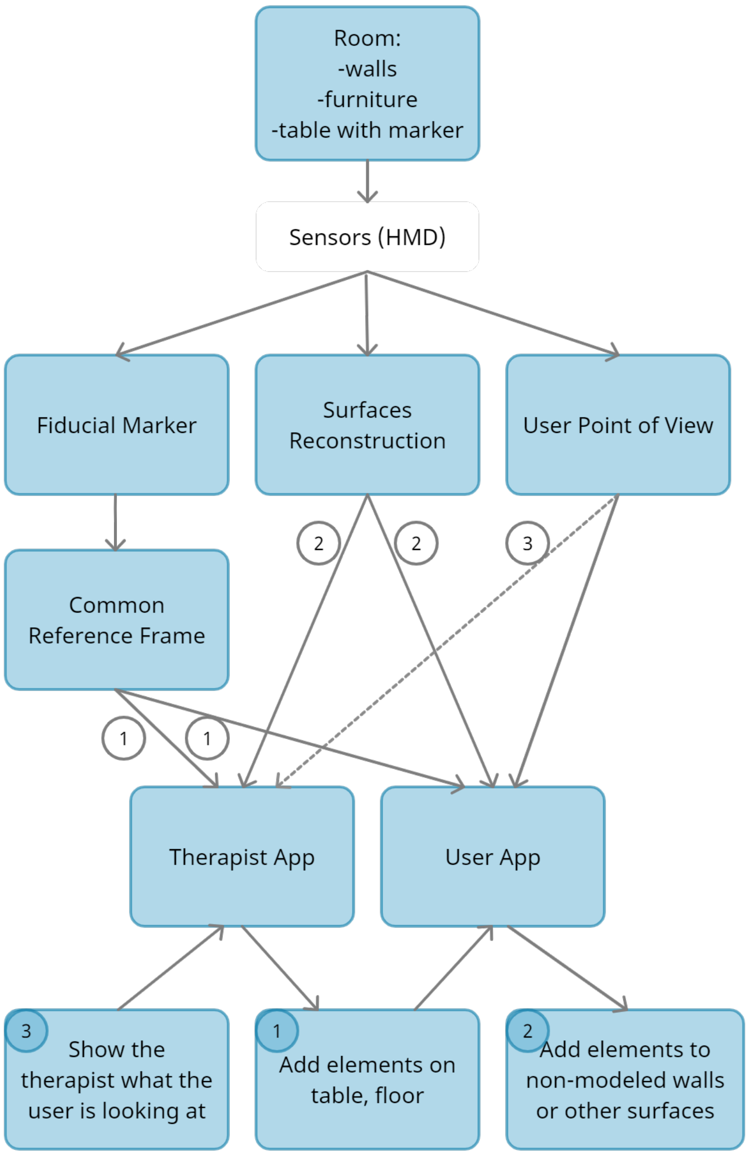

Figure 3.2: Flow diagram of the solution.

3.2 Materials and Development Tools

3.2.1 Materials

As we already saw in the proposal, the solution needs two devices: a computer and a Head Mounted Display (HMD), preferentially optimized to augmented reality.

Hololens 2

For implementing the proposed solution we need a device with the right technologies for our requirements. Since the solution will be applied in a therapeutic context, which requires an analysis of the patient's expressions, we aim to find a device that does not obstruct the user's face. To be comfortable, the device should be wireless and lightweight. Finally, the chosen HMD should have all the cameras and sensors necessary for the holograms to be realistic and to be able to implement the requirements defined in this work and in possible future follow-up work.

The target device chosen was Microsoft Hololens 2 (figure 3.3). These glasses have see-through holographic lenses with a display with a resolution of 2k 3:2 light engines, a holographic density of more than 2.5k light points per radian, and display optimization for 3D eye position. In terms of sensors, this device has 4 visible light cameras, 2 Infrared (IR) cameras, a Time-of-Flight (ToF) depth sensor, accelerometer, gyroscope, magnetometer, and a camera with a resolution of 1080p30 for video, speakers, and a microphone array with 5 channels. These sensors allow hand tracking, eye tracking, world-scale position tracking, and spatial mapping [35].

These features are extremely useful for the development of the proposed work. The see-through holographic lenses allow for increased realism and immersion of the models without increasing the computational burden to create and display them. The automatic optimization of the device for each user's eye position also increases the realism of the added virtual elements. Light, IR and depth cameras allow the detection and estimation of the position of objects such as markers, the patient's hands, along with others. Finally, to know the position and orientation of the user glasses in real-time we utilize several sensors such as the gyroscope, accelerometer, among others.

In addition to these features and sensors, this HMD was chosen due to the existence of available tools that ease the use of the Hololens 2 functionalities, allowing us to focus on the development of the application.

Figure 3.3: User interacting with AR via Hololens 2.

3.2.2 Development Tools

Mixed Reality Toolkit

As seen before, the Hololens 2 is a mixed reality device that is equipped with a variety of sensors and capabilities. To fully explore these capabilities and create mixed reality experiences, developers can use the Mixed Reality Toolkit (MRTK).

The MRTK is a toolkit designed to accelerate the development of cross-platform mixed reality applications. It provides a set of components and features that can be used to build and create mixed reality experiences. This toolkit is designed to simplify the development process and make it easier for developers to create high-quality mixed reality experiences for the Hololens 2. The development for this device can be made using directly MRTK or using a third-party platform commonly used in game development such as Unreal or Unity.

Unity

For the development of the application, there are two alternatives: implement directly using the toolkit presented earlier, or use a platform such as Unity or Unreal. Since the development platforms allow the implementation of the application to be accelerated, the direct use of the toolkit was discarded. Of the possible development platforms, Unity was chosen since this was the platform with which the lab members are most familiar.

Unity is a cross-platform game engine that allows the development of applications for a variety of desktop, mobile, console, and virtual reality platforms. This software disposes of several assets either closed and proprietary, or open source and royalty-free which supports different technologies from model importers, communications, or others. Examples of these assets are those provided by hardware vendors that support the development of applications for their specific platforms through some sort of hardware abstraction layer that hides the particular details. These assets, therefore, provide the necessary interfaces for the developer to extend them through their own assets and scripts. This platform was then chosen because of the abstraction it gives us by facilitating some work, such as model creation or interface design, allowing us to dedicate more time to programming the models and the desired capabilities of the application.

In our case, the goal of this work is to use realistic pre-made models to create a realistic environment, so the use of this platform facilitates our work since we can simply add these models to the work developed. Afterward, we create the necessary code for them to behave as desired.

4

Implementation

So far, we have been discussing our proposal for the dissertation, understanding the fit of this idea with the therapeutic context, and the problems to be solved. We will now analyze the implementation carried out during the thesis.

4.1 Overview

To understand the implementation, we must first analyze the structure of the application, in order to understand why we developed each component. Figure 4.1 shows the flow of the application. In this diagram, we can see the principal components that need to be developed. First, and once the patient and therapist have connected their devices, we need to establish the connection between the Hololens 2 and the therapist's computer. It is this connection that will later allow the exchange of information regarding the type, position, and rotation of the virtual models, among others. Then, we have to figure out how and where we can insert the virtual models so that they are presented in a realistic way. For this to happen, we will detect a QR code placed on the table where the therapy will take place, so its position is estimated and, consequently, we have the position of the therapy area. Finally, the therapist must be given the opportunity to add, remove or move the phobic elements in the position chosen by the therapist. With this in mind, we will now explain the implementation behind each one of these implementation steps.

4.2 Multiplatform communication

To allow communication between the Hololens 2 and the therapist's computer, we are using Unity Multiplayer Networking Netcode, a Unity networking library that allows faster

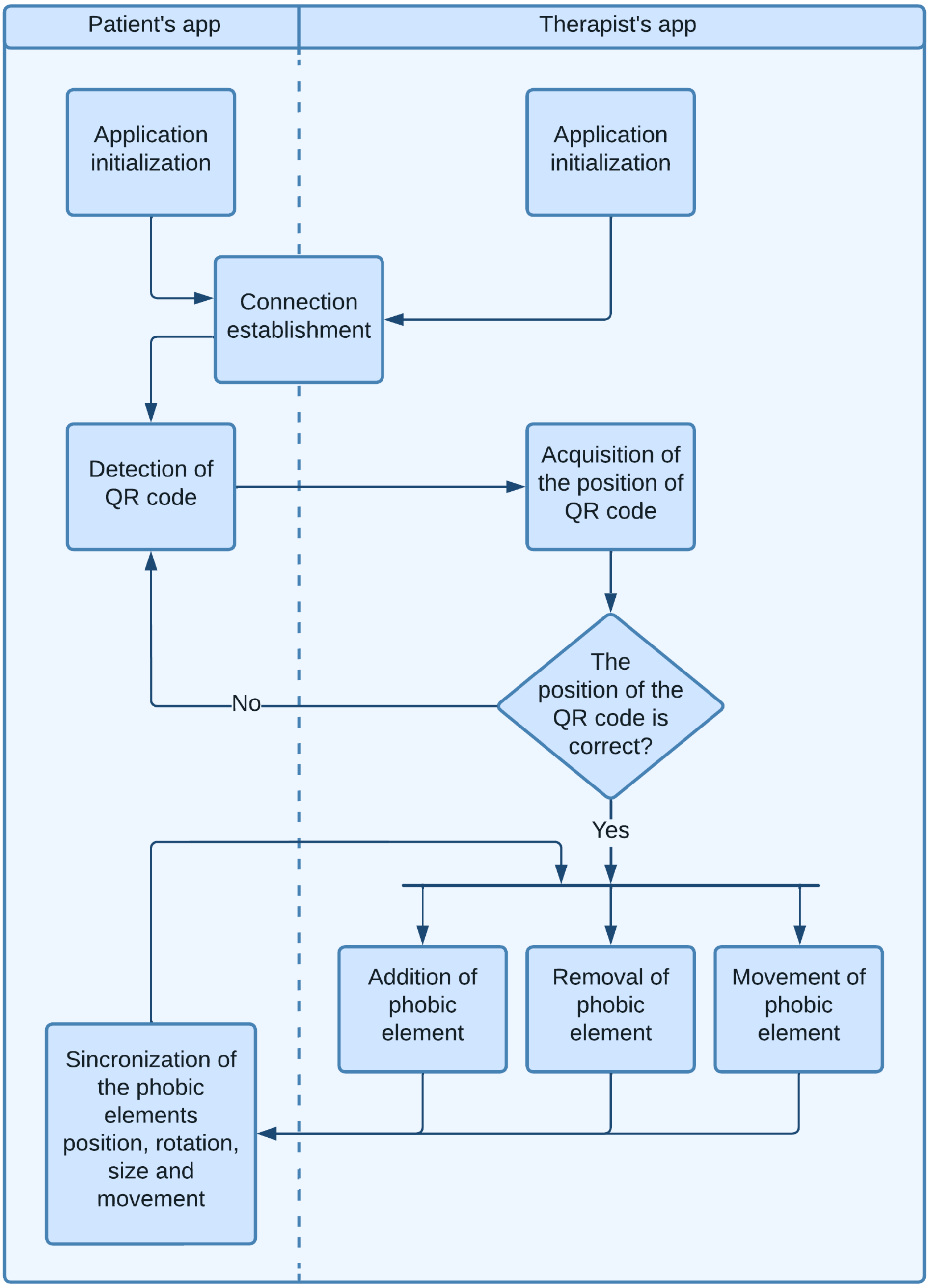

Figure 4.1: Overview of the implementation of the proposed system.

creation of networked applications [36]. This uses UNet transport that relies on User Datagram Protocol (UDP) transport layer to support the timed synchronization of the information between devices. The use of the UDP transport layer protocol is important because it is better suited for real-time data sharing, supports broadcasting, and takes up less bandwidth since fewer packets are needed for the same set of data than with Transmission Control Protocol (TCP) [37]. Given the nature of our problem, this is an adequate solution, since it is essential to have a real-time communication, even if there is no guarantee that all packets will reach their destination.

Network Discovery

To start the communication, we need to establish the connection. To create the connection we need the client to know the server's Internet Protocol (IP) address. To implement this IP address discovery we have three options: manually enter it, use a directory/game server, or perform a broadcast on the Local Area Network (LAN) network. The simplest option is the manual insertion of the IP address to be used in the connection, however, this requires the user to know the server's address. Since we are developing equipment for therapeutic use, which we intend to be used by anyone, even if they have little technological knowledge, this solution does not meet the proposed objectives. In this case, we want Zero-configuration networking (zeroconf), which is a set of technologies and protocols that allow devices to communicate with each other on a network without manual configuration. By eliminating the need for manual network configuration, zeroconf makes networking more accessible and user-friendly. In figure 4.2 we can see the three core technologies in which zeroconf is built. In this solution, we will focus on the automatic location of network services, which can be achieved with the two alternatives explained below. The second alternative would be to use a server where both the host (the therapist's computer) and the client (the Hololens 2) would connect and which would allow the elements to initiate the connection. However, this solution is more suitable for applications developed over the Internet. In our case, since the product is always intended to be used in a therapy context, that is to say, with the therapist and the patient in the same room, something so complicated is not necessary. We then have the last option, which consists of performing a broadcast to the LAN network, so the client can identify the IP address of the host and server. Due to usability and flexibility concerns and to enable highly flexible usage of the setup, without relying on directory/game servers that would need permanent

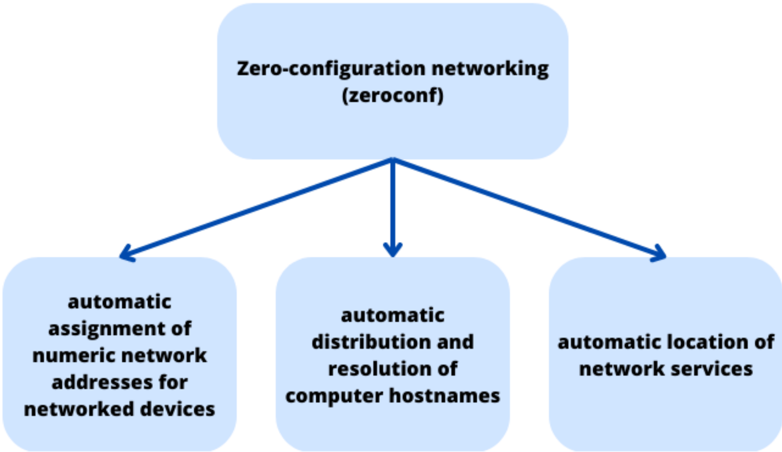

Figure 4.2: The three core technologies in which zeroconf is built.

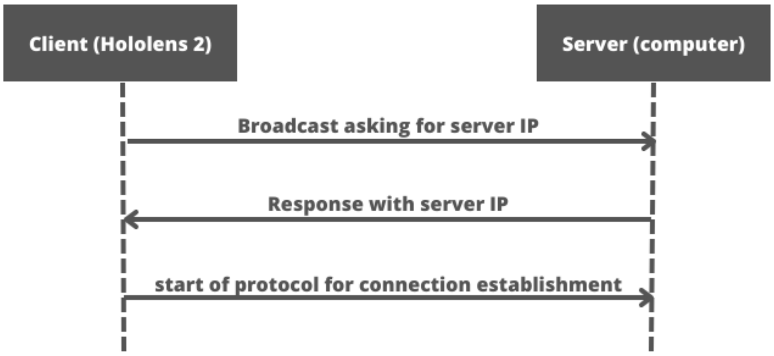

Figure 4.3: Diagram of the discovery of the server IP address.

internet connectivity, this was the option chosen.

Figure 4.3 shows a simple diagram that represents this discovery of the server IP address. This discovery step is based on an initial broadcast of a discovery message started by the AR device (the client), expecting to receive a response from the therapist's computer (the server). After this phase, the peers are mutually identified, and the connection is established supporting all the necessary communications used to place, move or destroy the phobic elements, parameter change, or other.

During the execution of the application, we have different types of communication. We may want to communicate only one value, like the position of the QR code, or keep several values synchronized, like the movements of the phobic elements. The Netcode allows us to do these various actions more optimally, through network variables, network objects, Remote Procedure Call (RPC), among others.

Network Object and Network Behaviour

In order to make each communication, any object used in the implementation must have a Network Object and a Network Behaviour component. The Network Object assigns to each object an identification (ID) used to associate the objects across the network. The Network Behaviour component is used to create the principal logic associated with the network. Both these components allow us to identify the components that, if desired and/or needed, are shared across the network.

Network Variables

A Network Variable is a simple way to synchronize a variable between server and clients without having to use RPC. To implement these, we start by defining the type of variable that we want to use, assigning the appropriate permissions, and specifying who can read and/or write those. Each time this variable changes, every client receives this updated value, including the new clients.

Network Transform

Each phobic element added during the application is an object whose position, rotation, and scale need to be synchronized between the server (the therapist) and the client (the patient). To do this synchronization we use the Network Transform component. For each position and rotation of the object, the values are shared between all the clients and server, which will apply this transformation in his objects, with interpolation to smoothly change the position.

Network Animator

At last, we need to implement and synchronize the animations of the phobic elements. This includes synchronizing the states, transitions, and properties of each object. The synchronization of all the elements of animation is made through RPC, which is sent every time a property changes.

All these components made it possible to create an application with communication. In the following sections, we will explain when these components were used.

4.3 Establishment of a common reference frame

To create a realistic therapy environment it is necessary that the elements that are built into it are also realistic, and for this to be achievable they must imitate real life as much as possible. In addition, the therapist's control of the therapy begins with controlling the real-world position of the virtual models. To make this happen, it is necessary to know the position of the patient and where the region where therapy will take place is located. For the sake of not overloading communication channels, it was decided not to share the image seen by the patient through Hololens with the therapist. However, it is necessary for the therapist to know what the patient is seeing and where to add the virtual elements. Therefore we can understand that some of the most important physical elements should have a virtual representation. As the Hololens device establishes its own coordinate frame, it is necessary to infer one transformation per physical object that will have a virtual counterpart representation in order to support the proper display of the virtual representations in the therapist's view.

Starting with a table, typically placed between the patient and the therapist, and where most of the interaction and exposure will be happening, a simple marker purposely placed on it can be enough support to the alignment of the digital model. In our case, a QR code is placed in the lower left corner of the table, so that the patient can look at it in the beginning, and then don't be affected by the presence of this marker. In fact, once detected, and given that the Hololens has a quite stable pose estimation, the relative transformation of the QR code is not expected to change even if the marker does not return to the camera field of view for a period of time. Therefore, we activate the QR code tracking in the beginning, and then deactivate it, to save computational resources and not affect the immersive experience of the user.

The detection of the marker is made using the cameras present in the Hololens 2. With the images captured, the QR code is detected and its pose estimation is consequently performed. Is then necessary to transform the QR code's real world coordinates into the Unity coordinate system. With this step done, we obtain the position and rotation of the QR code that we are going to use, and send these values to the therapist using Network variables. This information is used to place the table that the patient is seeing in the therapist's virtual environment, so that they can start placing the phobic elements. If the therapist is not satisfied with the position of the table (due to a misplacement of the QR code or a detection error), the detection can be repeated. After this mandatory step is performed, the positioning of the virtual objects in the real world is automatically done on the table, and the movement of the objects is also done on the table. This means that the virtual objects will never cross the table or be floating in the air, actions that would make the models unrealistic.

4.4 Controlling the phobic elements

Having established a connection between the patient and the therapist's application, and having detected the QR code that enables the establishment of a common reference frame, we are now ready to start adding the phobic elements used in the therapy. In figure 4.6 we can see an activity diagram with the addition and control of the phobic elements. We will now talk about each of these functionalities, and then explain how they are interrelated.

As stated before, we wanted a versatile therapy, that can be used for multiple problems and phobias. To accomplish this goal, we allow the use of multiple phobic elements. Due to the models available in the Unity Assets Store (the official store of assets to be used in Unity), we implemented the application with two animals: spiders and frogs, with one model for each animal. So, the therapist can choose the phobic element that is appropriate to each patient, according to their phobias. There is also the possibility of adding more than one type of phobic element at the same time. Although only two animal models are available at this moment, the application is made so that the addition of new models is simple, being enough for the programmer to configure the application for the existence of more models.

When we try to make a realistic model of an animal, we cannot make it without adding animation to the movement of the model. Animations enable the legs and arms of the animal model to keep up with the model's movement. Even when an animal is stationary, is not completely still. The frog, for example, has a crop that he moves even when is still. To make the models realistic and have an impact on patients, the therapist can choose if they want or not the animal to have an animation when it is not moving. When the animal is placed, it can also be moved to the front, sides, and can jump.

In the environment, the animals can have different sizes. The size of an animal can affect

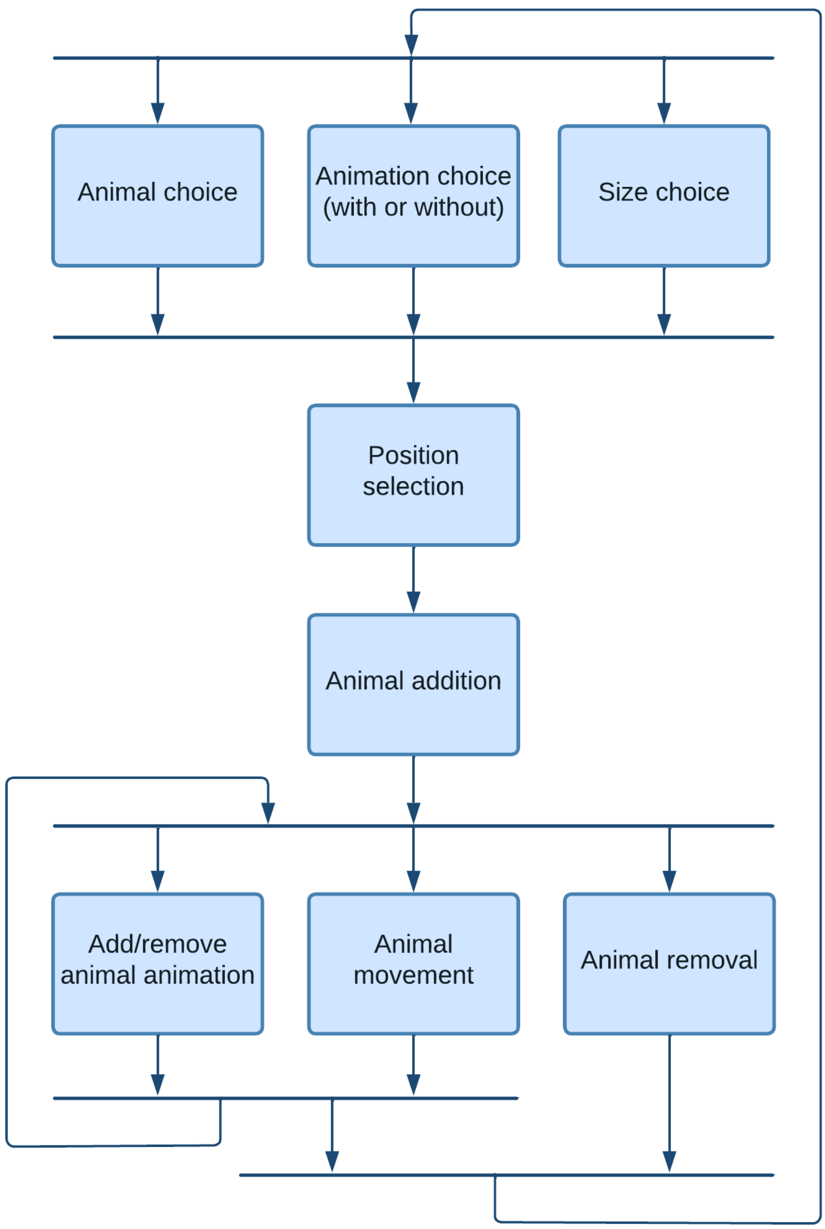

Figure 4.4: Addition and control of the phobic elements.

the patient more or less than expected, so the therapist can also decide what size the animals to be added should have. The therapist then has a slider that can be used to increase or decrease the models' size, between the pre-defined minimum and maximum size.

At last, it is important for the therapy that the doctor can choose the position of the phobic element, so that for an extreme case of phobia the animals can be farther, and with the improvement of the condition, the animals can be placed closer. In order to implement this idea, the therapist needs to observe the table where he wants to insert the phobic elements from a top view. This view, like others we will discuss later, is obtained simply by clicking a button. Once in this view, the choice of where to position the model is made by right-clicking on the desired position. That is, the therapist can see the virtual representation of the table seen by the patient, and uses it to precisely define the position of the models to be added. Similarly, to move or remove an animal, the therapist simply has to right-click on the phobic element that they intend to interact with, and then, choose the respective action.

To move a phobic element, after choosing the appropriate option, the doctor can control the virtual model using the keys on his keyboard. The up key allows the model to move forward, the right and left keys make the animal rotate to the right and left, respectively, and the down key enables the phobic element to jump forward. During this movement, the virtual animals are not static, they have an animation associated with them to make them more realistic.

To make the communication of all these elements between therapist and patient, the components Network Transform and Network Animation, as explained before, were used.

4.5 Animations and transitions

As mentioned earlier, for virtual models to be realistic, they have to be animated. We can consider that this animation is divided into three parts: the definition of the skeleton, the type of movements to be performed, and the activation of each action. Since creating a virtual model with animation for any animal is a time-consuming process that does not fall within the scope of this work, the models used were obtained from the Unity Asset Store. The following parameters were taken into account to choose them: the realism of the model, whether it is animated, and the percentage of people with phobias related to

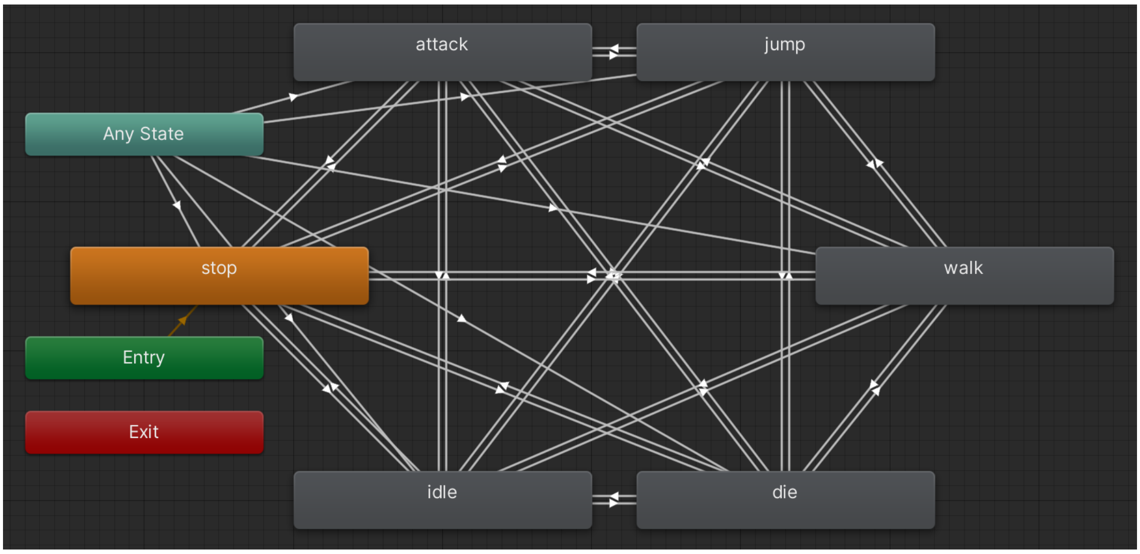

Figure 4.5: Diagram of the animations' transitions.

that animal. From this research came the two models currently used: the spider and the frog.

Once the models were chosen, we proceeded with the treatment of the animations. Since the models were not created by us, it is necessary to modify and adapt some animations to match what we intended. In the case of the frog model, it already had the following animations: walk forward, jump forward, turn left and right, and idle animation (for the situation explained above of the animal being stationary but not completely still). Therefore, it was only necessary to create a completely static animation and add it to the diagram of the animations' transitions.

The spider model used was not so complete, so it was necessary to add some animations. This one had walking forward, jumping, and idle animations. From these were created the stationary animation and adapted the walk animation for the rotation to the left and right. In addition, this model had no transition diagram of animations made, so we had to create one, shown in figure 4.5.

The animation transition diagram shown is a diagram where we link all the animations we want and where we define the animation change conditions. Initially, we define an initial animation, which will start with the introduction of the object into the scene. Next, we create the connections between all the animations, defining the animation change condition, which in our case are triggers. So, when we want to change the active animation, we simply activate the appropriate trigger. However, it is important to note that this diagram can have unused animations, as is the case since applications can undergo changes that make use of these animations.

Having created this diagram of transitions for each existing model, it remains for us to activate the animations when they are chosen by the therapist. As explained earlier, having done this work previously, we just need to activate the correct trigger, doing it also with the network animator so that the animation is shared with Hololens 2.

4.6 Extra features

As the application was developed and tested, the need arose to implement extra features to improve functionality.

4.6.1 Therapist's multiple views

As we discussed before, to not overload the communication channel, the therapist does not see everything that the patient sees, just the table where the phobic elements are added. However, depending on where the therapist is in relation to the patient, could be interesting having different views of the therapy and of the interaction of the patient with the phobic elements. Therefore, we decided to add these different views of the therapy, in order for the therapist to better observe and control the patient's response to the treatment. We then have the following views:

- **up view** - in this view the therapist sees the table from above. Adding the phobic elements and controlling them is done from this perspective.
- **front view** - in this view the therapist sees the table from the patient's point of view.
- **back view** - in this view the therapist sees the table as if facing the patient.
- **right view** - in this view the therapist sees the table as being to the right of the patient.
- **left view** - in this view the therapist sees the table as being to the left of the patient.

4.6.2 Patient's hands' detection

In the therapeutic context, it is essential for the therapist to be able to analyze the patient's reactions and interactions with the phobic elements in order to adjust the therapy. The chosen HMD already allows for the analysis of facial expressions. However, since the added animals are virtual, by simply observing the patient it is not possible to extract information about the position of the patient's hands in relation to the phobic elements present. To mitigate this problem, it is necessary to share the position of the patient's hands with the therapist. To do this, we started by detecting the position of the hands with the Hololens, taking from this detection the points corresponding to the palm and index finger. Later, the positions of these points were shared with the therapist using the Network Variable component discussed earlier. With the necessary positions shared, it is then possible to position the obtained hand models in the Unity Asset Store, making the position of the patient's hands visible to the therapist.

4.7 Interface

It is now time to present the interface developed in this work. This interface has two sides: the therapist and the patient side. The goal of the patient side is for the person to focus on the virtual models. This led us to create a simplified interface that only shows the essential elements, i.e., the phobic elements. Figure 4.6 presents the models of the animals currently used in the solution, and figure 4.7 offers the patient's view through the Hololens 2 of the two models currently present in our work: the spider and the frog. These models are two examples of what can be done with our application, since the framework is made to add more models easily.

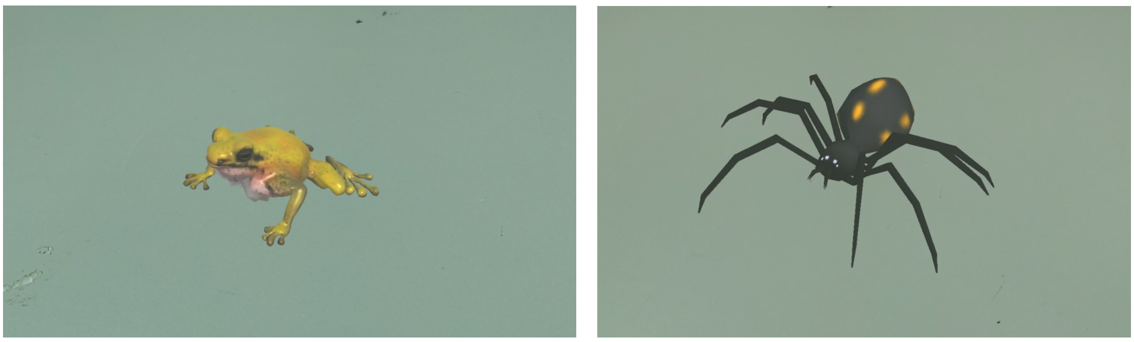

Figure 4.6: Model used in the application.

Figure 4.7: Patient view of the therapy (with the spider and frog models).

The objective of the therapist's side is to give the therapist all the necessary tools for this type of therapy in a simple and intuitive way. To achieve this, we focus on two aspects: giving all the necessary information clearly and straightforwardly, and keeping the design simple and intuitive. Figure 4.8 shows three initial instructions presented in the solution: the information that the patient's device is not connected, the necessity of reading the QR code, and the confirmation of the position of the QR code.

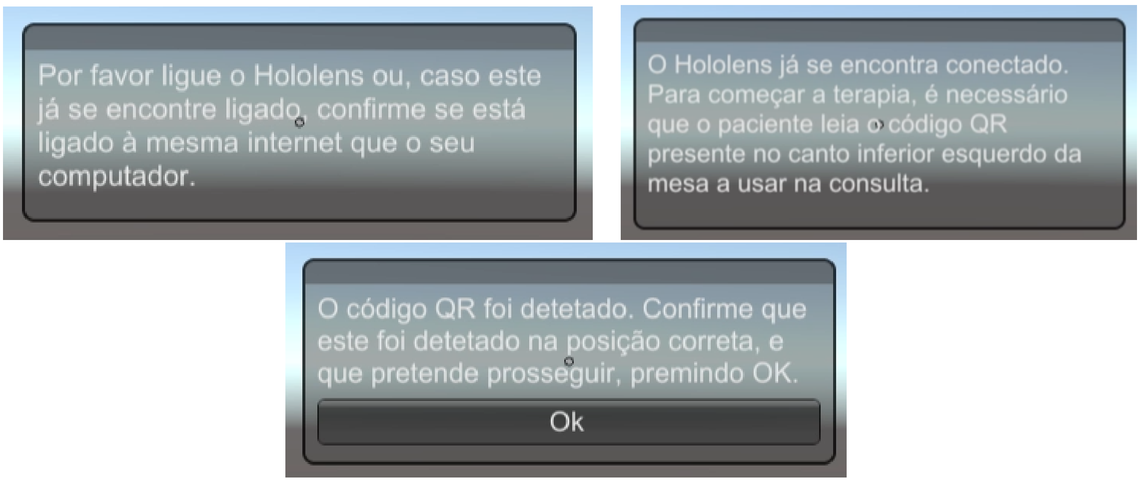

Figure 4.8: Three examples of instructions included in the application.

After this initial phase, the therapist is ready to start adding the phobic elements of his or her choice. To make this addition possible, it is necessary to choose which elements to add. So this choice is given to the user, and, as we will see, this can be changed at any time. In figure 4.9 we see the windows that allow the therapist to choose the initial phobic element and whether it is entirely stationary or not.

Figure 4.9: Initial choice of the phobic elements.

After this initial choice, the therapist is ready to add the phobic elements. These are added in the ”top view” shown in figure 4.10. In this image we can see that this view is divided into 4 parts:

- **1-** instructions on using the application and on adding virtual models;
- **2-** set of buttons that allow changing the therapist's observation perspective;
- **3-** set of options for viewing the virtual models (model type, animation choice, and size);
- **4-** virtual representation of the table used in therapy where the phobic elements will be added.

In this top view, we can also see the representation of the hands of the patient, as explained in subsection 4.6.2.

Figure 4.11: Options available to control the phobic elements.

As explained before, this top view allows the therapist to add phobic elements. To add these elements we need to choose the position desired with the mouse and then right-click it. If we want to animate or stop the animation, move or destroy a phobic element, we also

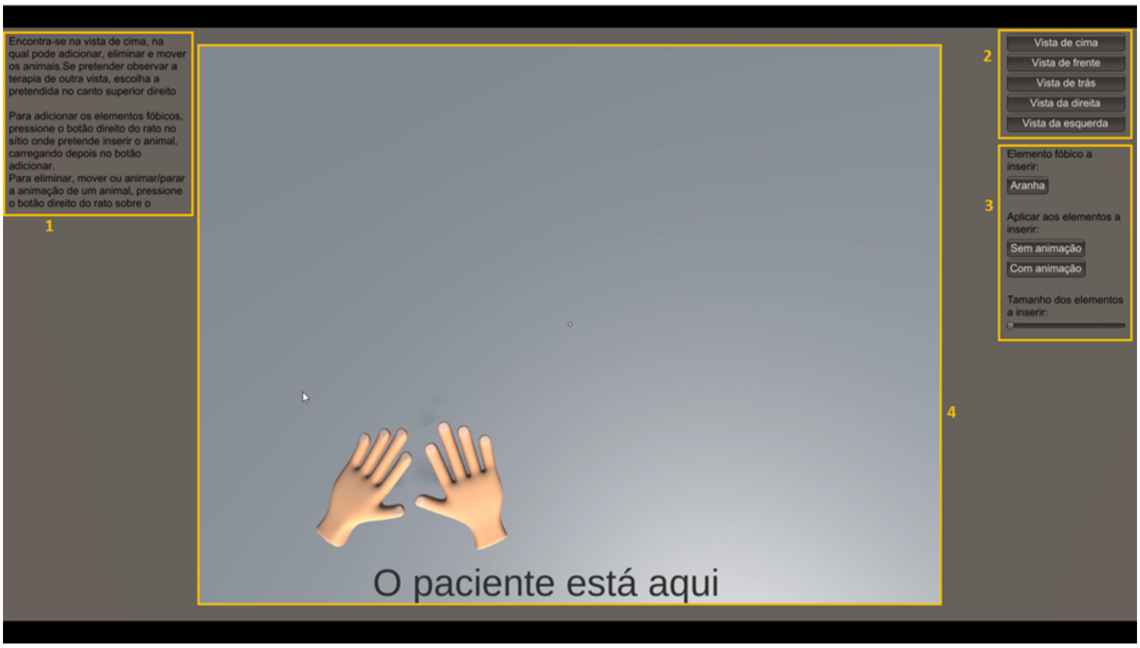

Figure 4.10: Therapist's top view of the therapy.

right-click with the mouse on top of the desired element. The options presented when we do this process are shown in figure 4.11.

In addition to the top view that allows the therapist to interact with the virtual elements, the therapist may also want to look at the therapy from other perspectives. Therefore, he or she has the possibility of choosing between the various views shown in figure 4.12: front, back, right, and left view. In this image, we can also see the representation of the patient's hands.

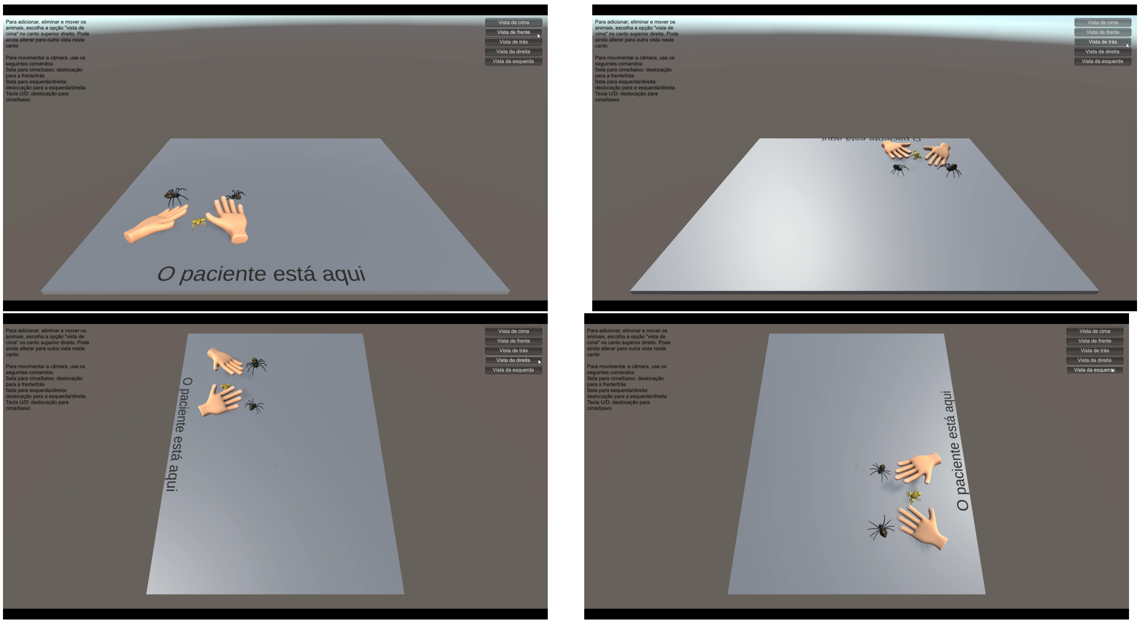

Figure 4.12: Views of the therapy (front, back, right, and left, respectively).

5

Results and Discussion

To evaluate the results, we decided to perform some studies with non-phobic people to assess the quality of the virtual models, the immersion of the AR application, and, on the other side, the usability of the therapist's app. For this, it was necessary to decide the type of trial and how many participants were necessary to perform each trial. For this decision, we followed the idea of Nielsen and Landauer [33], who stated that with only five users it is possible to find 85% of the problems in interactive systems. The authors argue that as the number of participants in a study increases, the amount of new information gained per participant decreases, making it less effective in identifying problems in the system. As a result, the authors recommend conducting several small studies with a limited number of users to identify potential problems early on in the development process, where in between studies the problems noticed are solved. In fact, early identification of potential problems typically brings an enormous advantage, as corrections tend to be simpler to apply at this phase, rather than over a final product.

With this in mind, a preliminary study was made during the development of the application with five users to assess the quality of the virtual models and their placement. The participants were two female and three male electrical and computer engineering students ranging from 21 to 30 years old. Each one saw 3 spiders placed on a table, with similar characteristics to the surrounding environment. The questions for this evaluation were chosen from the ”reality judgment and presence questionnaire” of Regenbrecht and Schubert [32] and are presented in the table 5.1.

In this study, question 6 was not made to the volunteers. The remaining questions were made, with answers that can vary between one and five. The results are presented in figure 5.1. This graph consists of a box, which represents the interquartile range, that is

Q1 - Was watching the virtual objects just as natural as watching the real world?	1= No	5= Yes
Q2 - Did you have the impression that the virtual objects belonged to the real object, or did they seem separate from it?	1= Separate objects	5= Same object
Q3 - Did you have the impression that you could have touched and grasped the virtual objects?	1= No	5= Yes
Q4 - Did the virtual objects appear to be on a screen, or did you have the impression that they were located in space?	1= On screen	5= In space
Q5 - Did you have the impression of seeing the virtual objects as merely flat images or as three-dimensional objects?	1= Flat images	5= Three-dimensional objects
Q6 - Did you pay attention at all to the difference between real and virtual objects?	1= I Saw no difference	5= The difference was clear
Q7 - Did you have to make an effort to recognize the virtual objects as being three-dimensional?	1= Yes an effort was needed	5= No effort was needed

Table 5.1: Questions of the “Reality judgment and presence questionnaire”

the interval that includes 50% of the data, and whiskers, which represent the dispersion of data outside the interquartile range. The average of the answers for each question is also marked with a dot.

From this study, we understand that the perception of the animal models highly depends on the person, since some participants did not find the spiders to be very realistic. Some participants also did not feel like the models were correctly placed in space and were tridimensional. Through the participants’ opinions, we also realize that the ambient light affects the quality of the perception of the virtual models. In reality, the holographic nature of the device is based on ”adding light”, and does not dispose of any means to block the light that comes from the objects. As a result, if the environment itself is too bright, the virtual elements tend to fade out.

With these problems in mind, we continued developing our work. Then, a second trial was made with ten males and one female divided into two groups: the first group of six people who interacted with frogs and the second group of five people who interacted with spiders. The participants have ages ranging from 20 to 30 years old. The conditions of light were

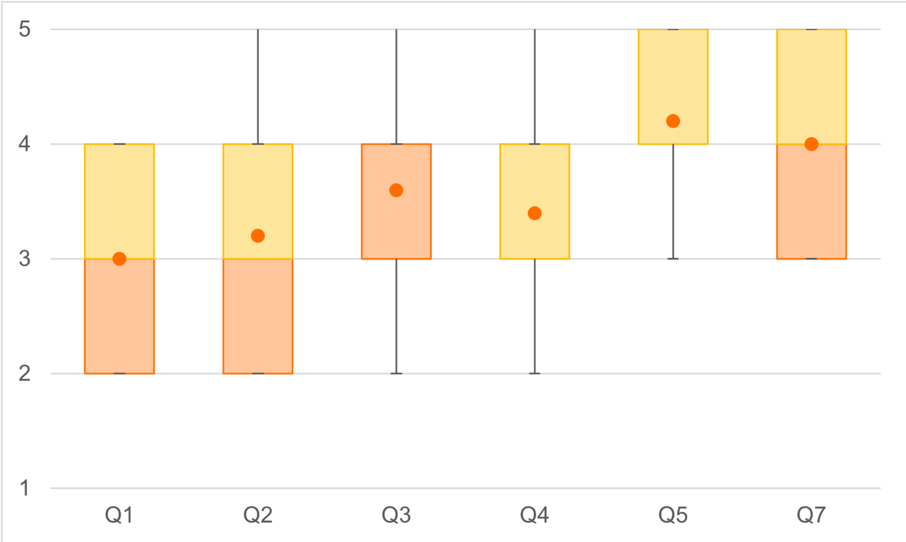

Figure 5.1: Boxplot of the answers to the first test.

similar, and all people saw three animals on a table (first static and then animated to look alive). The Hololens 2 was also calibrated before every test to adjust the view of the models to each user. Figures 5.2, 5.3 and 5.4 present the boxplot and the average of the answers to the questionnaires, for all of the answers, just the spiders and just the frogs, respectively.

From the observation of the answers to questions Q1 and Q2, we can see that we obtained almost the same responses for the different models. Despite the models not being perfect, the answers show that watching them is nearly as natural as watching the real world, in which the objects seem to belong. About Q3 and Q5, the distribution of the answers was slightly different, but the responses concentrated more on the value five, the maximum. That shows us that the objects were seen as being tridimensional and realistic, and not as flat images on a screen. Regarding questions Q4 and Q7, the volunteers who interacted with spiders gave a perfect score in Q4 and almost perfect in Q7, confirming that the models seem to be located in space. The volunteers who saw the frogs seem more aware of the screen used to create the holographic objects. However, from the positive results, we believe that this awareness does not affect the overall experience. When analyzing the values of question Q6, we understand that the difference between real objects and virtual objects was clearly perceived, however, the impact of this in the therapy has yet to be studied in a therapeutic context.

From the analysis of the results obtained, we can conclude that the models are realistic

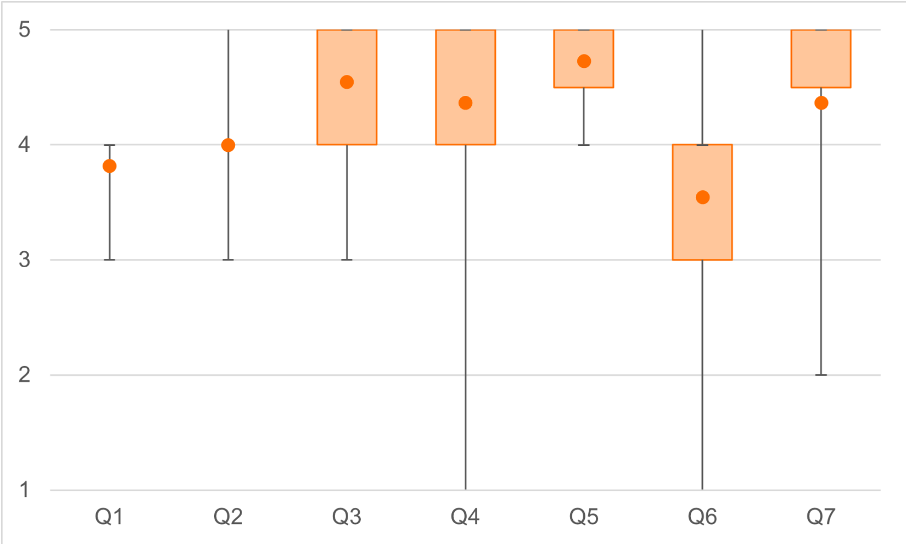

Figure 5.2: Boxplot of the answers to the second test.

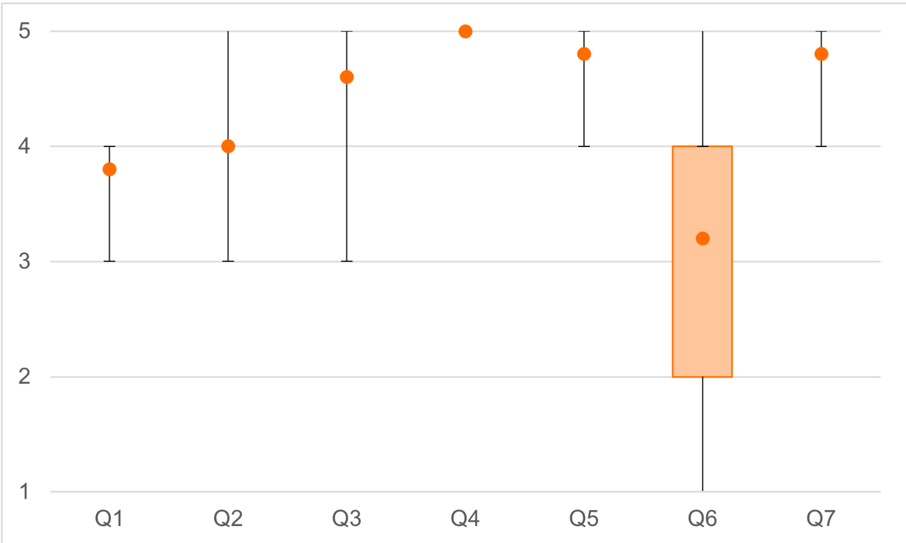

Figure 5.3: Boxplot with the responses to the second test of the participants who observed the spiders.

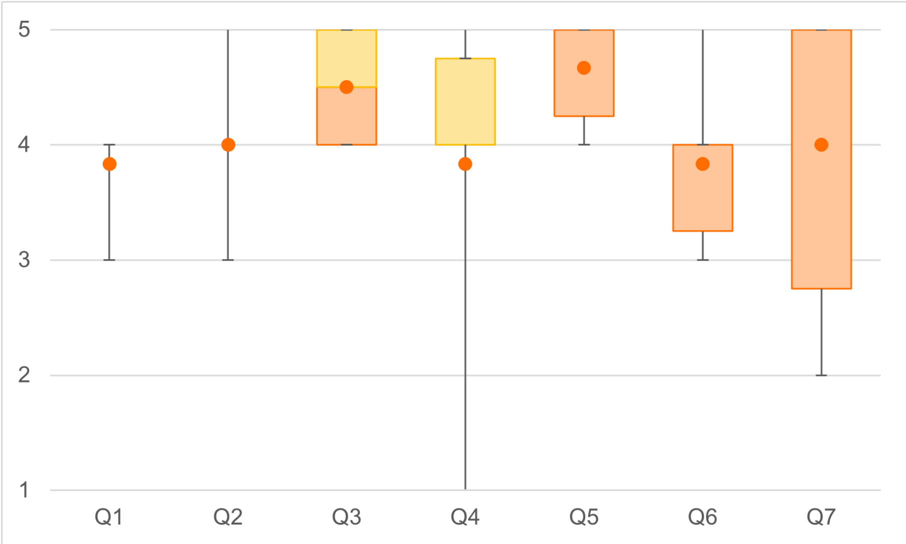

Figure 5.4: Boxplot with the responses to the second test of the participants who observed the frogs.

and produce the desired effect of appearing to be part of the environment. In addition to these answers, it was also observed that as soon as the volunteers saw the models, most of them had an instinctive reaction to try to touch or smash them, confirming the conclusions made.

With the two studies performed, we were able to evaluate the patient's side of the application, however, we are missing the evaluation of the therapist's side. For this, we conducted a third study involving 12 participants aged between 21 and 30 years old whose objective was to evaluate the therapist side of the application in relation to its usability, simplicity, and controllability. For this, each volunteer was given a list of tasks they had to perform for the test to be completed, after which they were asked to answer a questionnaire. The list of tasks to be performed was created based on the actions that a therapist has to perform in therapy and the functionalities provided. This list is as follows:

- Add at least one animal of each type (at least one spider and one frog)
- Move at least one animal
- Change the therapist's perspective on therapy at least once

To evaluate the work developed, the User Experience Questionnaire (UEQ) was chosen. This questionnaire presented by B. Laugwitz, T. Held and M. Schrepp [34] in 2008 allowed us to evaluate the user experience in the classical usability aspects (efficiency, perspicuity,

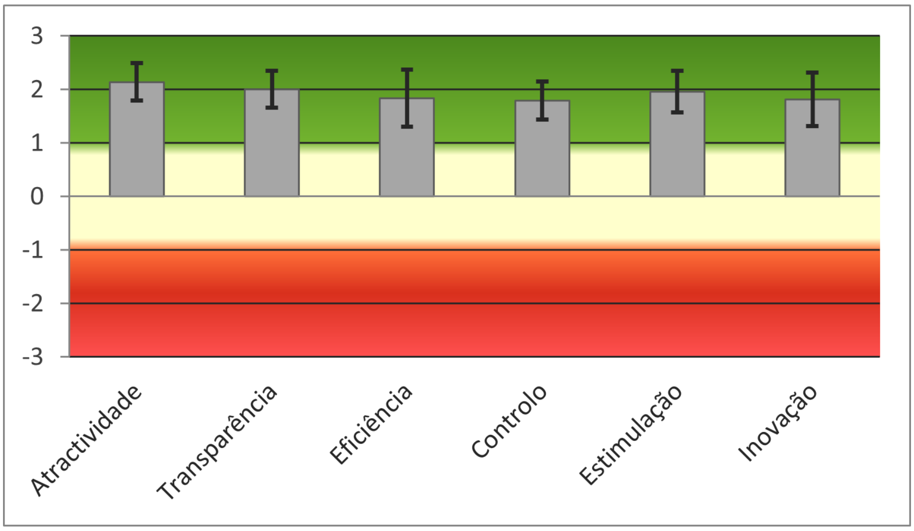

Figure 5.5: Mean and variance for all scales.

dependability) and user experience aspects (originality, stimulation). The questionnaire used is presented in Appendix. Figures 5.5, 5.6 and 5.7 show the results of this trial.

This questionnaire evaluates the results by scales, that is, the various questions are divided into 6 scales: attractiveness, transparency, efficiency, control, stimulation, and innovation. Figure 5.5 shows the mean and variance of the results divided by the scales. The authors of the questionnaire explain that below -0.8 the results are negative, between -0.8 and 0.8 are neutral, and above 0.8 are positive. The range of the scales is from -3 to 3, however, because of the variation in opinions and the tendency for people to avoid extreme answers, it is unlikely to find values above 2 or below -2. Therefore, with the lowest scale value obtained being 1.79 in the ”controlo”, we can conclude that, in general, the therapist side of the developed application meets the proposed requirements.

In figure 5.6 we see the distribution of responses per item, where we conclude once again that the results are positive. The worst responses obtained were in relation to the options ”lento/rápido” and ”imprevisível/previsível”, and these may have arisen due to different interpretations of the options. Users may consider the movement of the animals to be slow but the response of the application to be generally fast, or the movement to be unpredictable, even though the application’s commands are not, and take this movement into account when evaluating the application.

As the last graph of this study, we have a benchmark that allows to compare the quality of this application with those existing in the database associated with this study. By analyzing

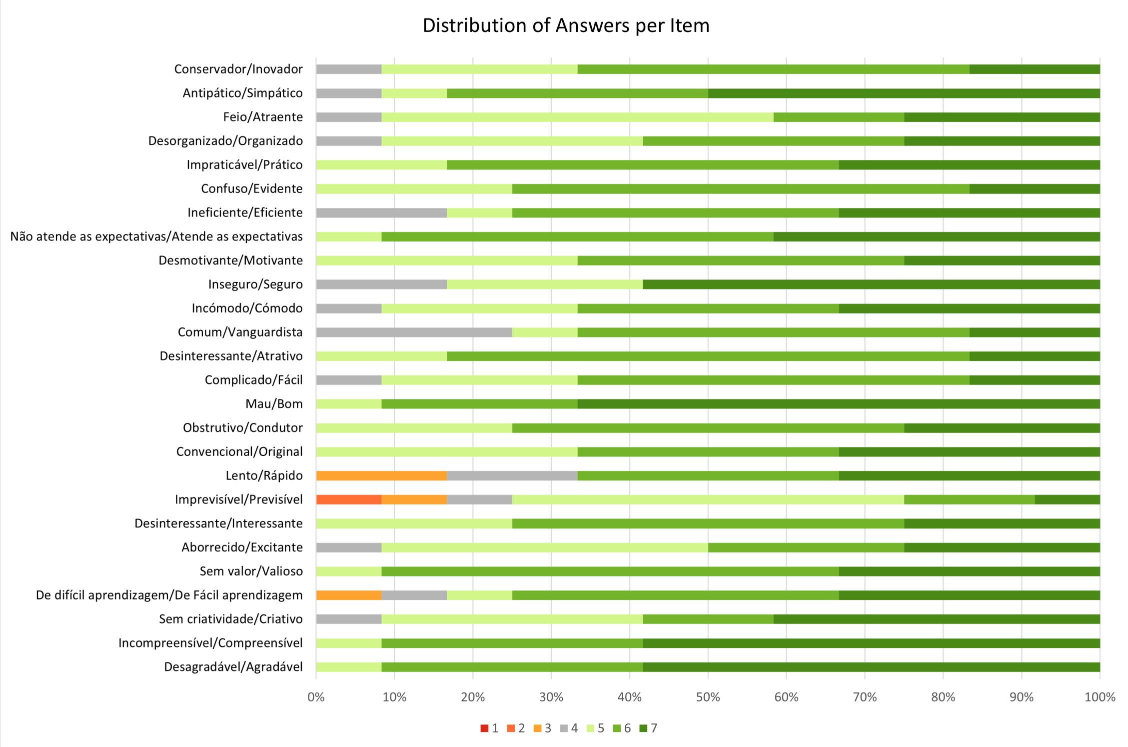

Figure 5.6: Answer distribution.

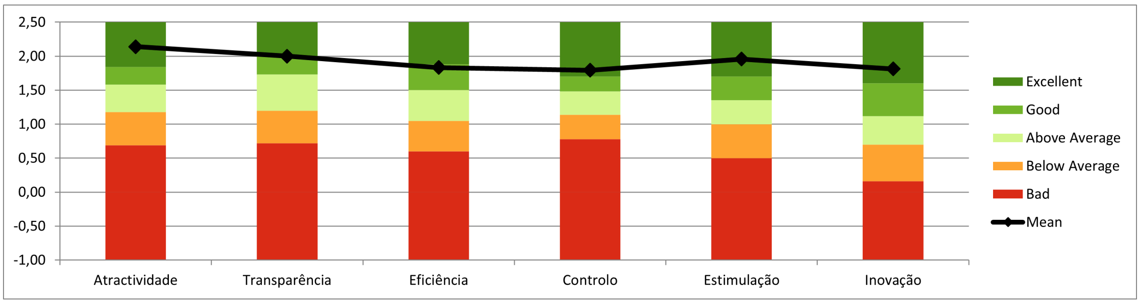

Figure 5.7: Benchmark for all scales.

the results, we realize that all scales have an excellent or good result, which once again confirms the success of the implementation of the objectives defined in relation to the therapist side of the application, namely its simplicity and control.

Finally, since this application is intended as therapeutic support, it is essential to be seen and evaluated by a therapist. This validation by a therapist was then carried out, where the idea behind the developed application was presented, the reason for the implementation choices made, and finally, the application was tested by the therapist, who could see both the patient and the therapist's side.

The feedback obtained by the therapist was very positive, validating the patient's simple side and the choice of HMD's, which allow not only the patient to see their surroundings, but the therapist to observe the patient's facial reactions. Regarding the therapist side of the application, its simplicity and usability were also validated as being what is needed to perform therapy. The possibility of using the familiar office environment was also highlighted as being very valuable, since it is important that the therapy is done in a familiar environment. From the presentation of the implemented solution also came the discussion of the most usual phobic elements in therapy that could be added in the future, such as cockroaches and ambulances. The idea of including sound in the application also came up, which would be completely controlled, just like the animals, and could even be added before the phobic elements, giving the sensation of proximity but not visibility of them. Although these suggestions were not implemented in this work, they were communicated to the colleague responsible for the continuation of the same, so they will be included in the future.

6

Conclusion and future work

The work presented in this dissertation was aimed at creating a therapeutic solution to replace *in vivo* exposure in the treatment of phobias. It was also important to create a simple and intuitive solution that would not be distracting or take the focus off the therapy and the therapist. With this in mind, we created a solution divided into two parts, the patient's part which uses augmented reality to present the phobic elements, and the therapist's side, which allows the use of a computer to control the entire therapy. It is even made to span multiple therapies, by being able to add more than one phobic element.

Because the application was split into two parts, the tests performed also had to be divided. The first tests confirmed that the virtual models presented in the HMD are realistic and look three-dimensional. The tests on the therapist's side, on the other hand, showed that the computer application was intuitive, attractive, and easy to use. Finally, this solution was presented to a therapist who found it useful, fit for purpose, and simple to use.

In addition to the studies done, it would be interesting to conduct a study on people with phobias to evaluate its real impact on its target audience. However, a study of this kind can take more than a year, since it has to be approved, it is necessary to find several people with this condition, and it takes a long time to cure, requiring several sessions. For all these reasons, this study was not carried out, however, the analysis of the test results and the therapist's opinion lead us to believe that the solution developed may be an added value for both therapists and patients.

This thesis presents the initial work on a solution with immense potential. However, in the course of a year, it is not possible to implement all features. Therefore, it would be interesting to add more phobic elements in the near future, since the application is already prepared for this, and add sound, as the therapist suggested. It would also be interesting to integrate the eye-tracking capabilities of Hololens 2 to make a real-time analysis of the patient's focus of attention and whether he looks at or avoids the phobic elements.

During the development of this work, a paper on it was written, accepted, and presented at *International Conference on Graphics and Interaction 2022 (ICGI'2022)*. The author's version of this paper is included in the Appendix of this document.

Bibliography

[1] R.C. Kessler, W. T. Chiu, O. Demler, K. R. Merikangas, E. E. Walters,“Prevalence, Severity, and Comorbidity of 12-Month DSM-IV Disorders in the National Comorbidity Survey Replication”, Archives of General Psychiatry, vol. 62, pp. 617-627, June 2005.

[2] U.S. Department of Health and Human Services,“Anxiety disorders”, National Institute of Mental Health, Last accessed 4 January 2021, 2018, Available online: https://www.nimh.nih.gov/health/topics/anxiety-disorders.

[3] S. Dattani, H. Ritchie, M. Roser,“Mental Health”, Our World in Data, 2021.

[4] Y. Choy, A. J. Fyer, J. D. Lipsitz, “Treatment of specific phobia in adults”, Clinical Psychology Review, vol. 27, pp. 266-286, 2007.

[5] N. Morina, H. Ijntema, K. Meyerbröker, P. M.G. Emmelkamp,“Can virtual reality exposure therapy gains be generalized to real-life? A meta-analysis of studies applying behavioral assessments”, Behaviour Research and Therapy, vol. 74, pp. 18-24, 2015.

[6] C. Suso-Ribera, et al.,“Virtual Reality, Augmented Reality, and In Vivo Exposure Therapy: A Preliminary Comparison of Treatment Efficacy in Small Animal Phobia”, Cyberpsychology, Behavior, and Social Networking, vol. 22, pp. 266-286, 2018.

[7] P.M.G. Emmelkamp, M. Krijn, A. M. Hulsbosch, S. de Vries, M. J. Schuemie, C. A. P. G. van der Mast,,“Virtual reality treatment versus exposure in vivo: a comparative evaluation in acrophobia”, Behaviour Research and Therapy, vol. 40, pp. 509-516, 2002.

[8] T. D. Parsons, A. A. Rizzo,“Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis”, Journal of Behavior Therapy and Experimental Psychiatry, vol. 39, pp. 250-261, 2008.

[9] I. C. Giglioli, F. Pallavicini, E. Pedroli, S. Serino, G. Riva, “Augmented Reality: A Brand New Challenge for the Assessment and Treatment of Psychological Disorders”, Computational and Mathematical Methods in Medicine, pp. 1-12, 2015

[10] N. A. J. De Witte, S. Scheveneels, R. Sels, G. Debard, D. Hermans, T. Van Daele, “Augmenting Exposure Therapy: Mobile Augmented Reality for Specific Phobia”, Frontiers in Virtual Reality, vol. 1, pp. 8, 2020.

[11] M. Wrzesien, J. Burkhardt, M. Alcañiz, C. Botella, “How Technology Influences the Therapeutic Process: A Comparative Field Evaluation of Augmented Reality and In Vivo Exposure Therapy for Phobia of Small Animals”, Springer Berlin Heidelberg, pp.523-540, 2011.

[12] C. Botella, W. A. Pérez-Ara, J. Bretón-López, S. Quero, A. García-Palacios, R. M. Baños, “In Vivo versus Augmented Reality Exposure in the Treatment of Small Animal Phobia: A Randomized Controlled Trial”, PLOS ONE, vol. 11, pp. 1-22, February 2016.

[13] C. Botella, J. Fernández-Álvarez, V. Guillén, A. García-Palacios, R. Baños, “Recent Progress in Virtual Reality Exposure Therapy for Phobias: A Systematic Review”, Current Psychiatry Reports, vol. 19, pp. 1-12, 2017.

[14] D. Boeldt, E. McMahon, M. McFaul, W. Greenleaf, “Using Virtual Reality Exposure Therapy to Enhance Treatment of Anxiety Disorders: Identifying Areas of Clinical Adoption and Potential Obstacles”, Frontiers in Psychiatry, vol. 10, pp. 773, 2019.

[15] A. Garcia-Palacios, C. Botella, H. Hoffman, S. Fabregat, “Comparing Acceptance and Refusal Rates of Virtual Reality Exposure vs. In Vivo Exposure by Patients with Specific Phobias”, CyberPsychology & Behavior, vol. 10, pp. 722-724, 2007.

[16] B. O. Olatunji, B. J. Deacon, J. S. Abramowitz,“The Cruelest Cure? Ethical Issues in the Implementation of Exposure-Based Treatments”, Cognitive and Behavioral Practice, vol. 16, pp. 172-180, 2009.

[17] J. Szymanski, W. O'Donohue,"Fear of Spiders Questionnaire", Journal of Behavior Therapy and Experimental Psychiatry, vol. 26, pp. 31-34, 1995.

[18] R. M. Baños, C. Botella, M. Alcañiz, V. Liaño, B. Guerrero, B. Rey, "Immersion and Emotion: Their Impact on the Sense of Presence", CyberPsychology & Behavior, vol. 7, pp. 734-741, 2006.

[19] A. Bangor, P. T. Kortum, J. T. Miller,"An Empirical Evaluation of the System Usability Scale", International Journal of Human–Computer Interaction, vol. 24, pp. 574-594, 2008.

[20] B. O. Rothbaum, L. F. Hodges, R. Kooper, D. Opdyke, J. S. Williford, M. North, "Effectiveness of computer-generated (virtual reality) graded exposure in the treatment of acrophobia." American Journal of Psychiatry, vol. 152, pp. 626-628, April 1995.

[21] S. Côté, S. Bouchard, "Documenting the efficacy of virtual reality exposure with psychophysiological and information processing measures." Applied Psychophysiology and Biofeedback, vol. 30, pp.217-232, September 2005.

[22] C. T. Pitti, et al., "The combined use of virtual reality exposure in the treatment of agoraphobia.", Actas Esp Psiquiatr, vol. 43, pp. 133-141, July-August 2015.

[23] M. C. Juan, M. Alcaniz, C. Monserrat, C. Botella, R. M. Baños and B. Guerrero, "Using augmented reality to treat phobias," IEEE Computer Graphics and Applications, vol. 25, pp. 31-37, November-December 2005.

[24] B. Patrão, P. Menezes, N. Gonçalves, "Augmented Shared Spaces: An Application for Exposure Psychotherapy.", International Journal of Online and Biomedical Engineering (iJOE), vol. 16, pp. 43-50, 2020

[25] J. Bretón-López, S. Quero, C. Botella, A. García-Palacios, R. M. Baños, M. Alcañiz, "An Augmented Reality System Validation for the Treatment of Cockroach Phobia." Cyberpsychology, Behavior, and Social Networking, vol. 13, pp. 705-710, December 2010.

[26] M. Wrzesien, M. Alcañiz, C. Botella, J. M. Burkhardt, J. Bretón-López, M. Ortega , D. B. Brotons, "The therapeutic lamp: treating small-animal phobias." IEEE Comput Graph Appl, vol. 33, pp. 80-86, January-February 2013.

[27] F. N. Tripi, M. Selzer, J. A. Biondi, M. Larrea, "'Virtual and Augmented Reality Therapy Framework for Phobia Treatment", XXIV Congreso Argentino de Ciencias de la Computación (La Plata, 2018), 2018.

[28] S. Corbett-Davies, A. Dünser, A. Clark, "An interactive Augmented Reality system for exposure treatment", 2012 IEEE International Symposium on Mixed and Augmented Reality - Arts, Media, and Humanities (ISMAR-AMH), 2012.

[29] E. Toma, O. Balan, C. Lambru, A. D. B. Moldoveanu, F. Moldoveanu, "Ophiophobia 3D – a Game for Treating Fear of Snakes", 2020 IEEE 10th International Conference on Intelligent Systems (IS), pp.205-210, 2020.

[30] A. Plasencia, M. Escobar-Sánchez, "Interactive multiplatform software to mitigate arachnophobia through augmented reality", Simposio Iberoamericano en Programación Informática (Ibero-American Symposium on Computer Programming), pp. 108–124, 2018.

[31] F. Fatharany, R. R. Hariadi, D. Herumurti, A. Yuniarti, "Augmented reality application for cockroach phobia therapy using everyday objects as marker substitute", 2016 International Conference on Information & Communication Technology and Systems (ICTS), pp. 49-52, 2016.

[32] H. Regenbrecht, T. Schubert, "Measuring Presence in Augmented Reality Environments: Design and a First Test of a Questionnaire", arXiv, 2021.

[33] J. Nielsen, T. K. Landauer, "A mathematical model of the finding of usability problems.", In Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems (CHI '93), Association for Computing Machinery, pp. 206-213, 1993.

[34] B. Laugwitz, T. Held and M. Schrepp, "Construction and Evaluation of a User Experience Questionnaire", USAB 2008, vol. 5298, pp. 63-76, November 2008.

[35] Microsoft,"Hololens 2", Last accessed 23 November 2022, Available online: https://www.microsoft.com/en-us/hololens/hardware.

[36] Unity,"About Netcode for GameObjects", Last accessed 15 November 2022, Available online: https://docs-multiplayer.unity3d.com/netcode/current/about/index.html.

[37] Geeks for Geeks,"Differences between TCP and UDP", Last accessed 15 November 2022, Available online: https://www.geeksforgeeks.org/differences-between-tcp-and-udp/.

Appendix

- User Experience Questionnaire - Portuguese version
- ICGI 2022 - Full Paper

Por favor dê-nos a sua opinião.

A fim de avaliar o produto, por favor preencha o seguinte questionário. É constituído por pares de opostos relativos às propriedades que o produto possa ter. As graduações entre os opostos são representadas por círculos. Ao marcar um dos círculos, você pode expressar sua opinião sobre um conceito.

Exemplo:

Atraente	○	●	○	○	○	○	○	**Feio**

Esta resposta significa que avalia o produto mais **atraente** do que **feio**.

Marque a sua resposta da forma mais espontânea possível. É importante que não pense demasiado na resposta porque a sua avaliação imediata é que é importante.

Por favor, assinale sempre uma resposta, mesmo que não tenha certezas sobre um par de termos ou que os termos não se enquadrem com o produto.

Não há respostas "certas" ou respostas "erradas". A sua opinião pessoal é que conta!

Por favor, dê-nos a sua avaliação atual do produto em causa.

Por favor, marque apenas um círculo por linha.

	1	2	3	4	5	6	7	
Desagradável	○	○	○	○	○	○	○	Agradável
Incompreensível	○	○	○	○	○	○	○	Compreensível
Criativo	○	○	○	○	○	○	○	Sem criatividade
De Fácil aprendizagem	○	○	○	○	○	○	○	De difícil aprendizagem
Valioso	○	○	○	○	○	○	○	Sem valor
Aborrecido	○	○	○	○	○	○	○	Excitante
Desinteressante	○	○	○	○	○	○	○	Interessante
Imprevisível	○	○	○	○	○	○	○	Previsível
Rápido	○	○	○	○	○	○	○	Lento
Original	○	○	○	○	○	○	○	Convencional
Obstrutivo	○	○	○	○	○	○	○	Condutor
Bom	○	○	○	○	○	○	○	Mau
Complicado	○	○	○	○	○	○	○	Fácil
Desinteressante	○	○	○	○	○	○	○	Atrativo
Comum	○	○	○	○	○	○	○	Vanguardista
Incómodo	○	○	○	○	○	○	○	Cómodo
Seguro	○	○	○	○	○	○	○	Inseguro
Motivante	○	○	○	○	○	○	○	Desmotivante
Atende as expectativas	○	○	○	○	○	○	○	Não atende as expectativas
Ineficiente	○	○	○	○	○	○	○	Eficiente
Evidente	○	○	○	○	○	○	○	Confuso
Impraticável	○	○	○	○	○	○	○	Prático
Organizado	○	○	○	○	○	○	○	Desorganizado
Atraente	○	○	○	○	○	○	○	Feio
Simpático	○	○	○	○	○	○	○	Antipático
Conservador	○	○	○	○	○	○	○	Inovador

IS3TA - A mixed reality tool for exposure-based therapies

Marta Nunes[1,2] and Paulo Menezes[1,3]

[1]Dept. Electrical and Computer Engineering, University of Coimbra
[2]marta.nunes@isr.uc.pt
[3]PauloMenezes@isr.uc.pt

Abstract

In vivo exposure is the most common treatment for phobias, although it has several drawbacks that can be mitigated by adopting technological alternatives such as virtual reality or augmented reality. Augmented reality provides some advantages over virtual reality, including fewer modelling costs and a higher level of realism. As a result, the goal is to develop an alternate treatment to exposure *in vivo* using augmented reality with Hololens 2. The proposed approach allows the patient to interact with the phobic elements while simultaneously giving the psychologist complete control over them, allowing each person to have a unique and personalized experience based on their phobias. Implementation and preliminary analysis results are presented.

Augmented Reality, Immersive Systems, Phobias, Anxiety, Exposure therapy

1 Introduction

Anxiety affects 18.1% of adults, making it one of the most common mental diseases [1]. When it reaches excessive levels, it causes agitation, muscle tension, and difficulty concentrating, among other things. Some anxiety disorders are associated with phobias, which are aversions or fears of specific items or circumstances that cause significant discomfort to the individual who comes into contact with them [2] [3].

There are various treatments for these diseases, and while no single treatment works for all phobias, *in vivo* exposure is the most robust treatment applicable to most of them. In this type of treatment, the patient directly confronts his phobic stimulus to overcome it [4]. However, this therapy has many problems associated with it, such as low acceptance and high dropout rates [14, 16], complications related to the presence of the phobic element (especially if it is an animal) [11, 12, 14], the impossibility of performing this therapy in the usual office environment [7], and the therapist's inability to control the phobic elements.

To overcome the drawbacks of *in vivo* exposure, technological solutions have been introduced in therapies using virtual and augmented reality. As virtual reality (VR) creates simulations of the real-life feared environment/elements, several studies have found that using it can be just as successful as using *in vivo* exposure [4–8, 13, 14]. Being relatively new in this application field, through augmented reality (AR) the patient may see the real world with phobic elements added. Despite being newer, this treatment method yields positive results comparable to *in vivo* exposure [6, 9–12]. Both treatments minimize the problems discussed above and may allow therapists to have complete control of the phobic elements, allowing gradual and tailored exposures based on patients' anxieties [5, 9–11, 14]. Plus, technological solutions allow a wider variety of exercises and scenarios offered to the patient [11, 12].

In the case of exposure therapy applications, augmented reality has some advantages over virtual reality, which goes from less demanding modelling needs, to improved user experiences. In fact, in augmented reality, the environment does not need modelling, and as such, it has lower programming and modelling requirements [6, 11, 12]. On the other side, the integration of virtual items in the actual world contributes to a more natural integration and acceptance from the user's point of view, also contributing to a stronger sense of presence [9, 11].

However, both virtual and augmented reality demand caution in their application, as these aim at producing experiences that convey sensations comparable to those found in the real world. If successful they produce realistic impressions which may be in turn susceptible to inducing emotional involvement, discomfort levels, and even anxiety [10].

We can better understand which applications are best for virtual or augmented reality by simply considering that the former "transports the user to another virtual place". In contrast, the latter brings virtual elements to the user space, or eventually modifies some of the existing ones. From this, it is simple to understand that virtual reality is particularly appropriate for cases such as claustrophobia and acrophobia. On the other hand, augmented reality can be very effective for animal phobias and other elements that can be integrated naturally into the user's surroundings and familiar environment, without the risk of cybersickness.

In line with the above, this paper proposes an alternative to *in vivo* exposure that makes use of augmented reality. This solution consists of a new tool that offers the possibility of presenting a plethora of stimuli/phobic elements in a fully controlled way, but still keeping the therapist present, aware and in control of the exposure session. This represents a step forward with respect to many other VR/AR based approaches that commonly do not enable the therapist to be a full player in control at real-time stimulation level. In our proposal, the therapist may not only tailor each session but also add or remove elements according to the observed reactions of the patients.

The remainder of this paper is structured as follows: (1) some of the most inspiring related works, (2) the proposed concept and approach, (3) the implementation and results, and the concluding remarks.

2 Related Work

2.1 Virtual Reality

The first alternatives to *in vivo* exposure have emerged using virtual reality. The work of Rothbaum et al. [20] (1995) was one of the first to present an experiment with virtual reality using a headset. It proposed a treatment for acrophobia using various virtual environments: various bridges, outdoor balconies at different heights, and one elevator simulation. Although there were no treatment comparison group and the subjects were not patients seeking treatment, the study successfully showed that virtual reality could be used to treat anxiety disorders.

In Côté and Bouchard's work [21] (2005), the authors focus on treating arachnophobia with a VR environment consisting of 2 apartments with many rooms with different levels. There are few static spiders in the lower levels, while the spiders were bigger and moving in the most difficult levels. The study showed improvement in participants' phobia symptoms.

Pitti et al. [22] (2015) presented an alternative treatment to agoraphobia with VR, using polarized glasses and projecting the image on a special screen. In this experience, there were various scenarios: a square and street, an airport building and plane, a bank office, an elevator and underground car park, a beach, a highway, and a cableway. This study concludes that the patients that received combined treatment (VR and psycho drugs) improved significantly more than the others. However, the dropout rates with VR and *in vivo* exposure were similar.

There are several other examples of the use of virtual reality to this end, but with the appearance of usable augmented reality headsets, these have lately received increasing interest in particular due to the fact that it enables to explore the familiarity of the local environment to avoid unnecessary anxiety induced just by the unfamiliar virtual environments themselves.

2.2 Augmented Reality

Juan et al. [23] (2005) presented one of the first attempts to use AR as an alternative to *in vivo* exposure, using a VR HMD and a camera connected to a computer. This work, although exploratory, already supported the idea of having the therapist watch the exposure scenario on the computer monitor. The virtual elements used were spiders (three types: small, medium, and large) and cockroaches (one type), and these could move, be static, or even dead, if the patient killed them. The interaction provided to the therapist was based on markers, controlling this way where the objects would appear, controlling how many animals appear, their size, and movements. The patients to "kill" the animals also needed a specific marker.

Bretón-López et al. [25] (2010) developed an experience using an AR system to treat cockroach phobia. The system ran on a PC, and similarly, the video was captured by a camera attached to an HMD and connected to the PC via USB. The therapist could control the cockroaches' number, movement, and sizes. The experience reported good results, however, the intricate cabled connection caused discomfort and tiredness.

The work of Wrzesien et al. [11] (2011) and Botella et al. [12] (2016) also aimed to treat spider and cockroach phobias using AR solutions. The patients could observe and interact with the scene through HMDs, and the therapists controlled the phobic triggers' sizes, number, and movement. Despite the success of the therapy, [11] it is mentioned that the systems have functions (like the non-optimal user-interface dialogue) that are not useful or not optimal, and [12] tiredness and dizziness were reported due to the HMD.

De Witte et al. [10] (2020) focus on ensuring the validity of using an AR application for smartphone or tablet to induce fear and anxiety, an important prerequisite for effective treatments with AR. The user could choose between various animals (rats, pigeons, snakes, cockroaches, spiders, frogs, and dogs). The study shows that the application induces fear and anxiety, even though the use of a smartphone or a tablet also reduces the realism and perception of the virtual elements, and makes it difficult to interact with these elements.

Patrão, Menezes and Gonçalves [24] (2020) presented an AR experiment with spiders. The application uses tablets, one for the therapist and another for the patient, and both can control and add elements to the environment. The study reveals good usability and interaction, yet, the use of tablets reduces the perception of the virtual elements and makes it difficult the interaction with them.

Wrzesien et al. [26] (2013) approach AR therapy with a new device, the Therapeutic Lamp (TL), which projects the elements on a surface. This experiment focuses on cockroach and spider phobias, with three types of each animal, and the therapist controls the number, size, and movement of the elements. The experience shows that the TL is an effective tool for small animal phobia treatment, however, the fact that it uses projection instead of holograms can affect the sense of presence and realism.

After reviewing these articles, we realized that in many of the approaches for therapy with AR, the devices used are also not always the most appropriate, as tablets, smartphones, and TL are likely to affect the realism of the therapy elements and the sense of presence. Previously used HMDs can become uncomfortable and cumbersome due to their weight and wires. Our alternative, presented below, is therefore intended to overcome these problems.

3 Proposal

As with other papers, the goal of this proposal focuses on developing a passive augmented reality solution to be used in a clinical setting to replace *in vivo* exposure. This resulted in the development of an *Immersive System for Sensory Stimulation and Therapeutic Applications (IS3TA)* that is expected to enable the therapist to bring into the office space virtual versions of elements associated with specific phobias of each patient. This system should be integrated into a typical therapeutic office space, where a table can be used in front of the patient, on which the exposure elements will appear, and with which the patient will be asked to approach (when deemed appropriate by the therapist), or even interact. This approach may also allow us to explore the surrounding space, floor, walls, and ceiling to introduce these types of elements. For example, it is more natural a dog to appear on the floor than on a table, whereas a spider might be on the wall or on the ceiling.

3.1 Putting the therapist in control

One of the innovations introduced in this approach is the possibility that the therapist can control the whole exhibit through a computer, selecting the elements that should appear, their location and characteristics, etc. Similarly, they can be removed if it is found that there is some overexposure effect. It will also be possible to choose the most appropriate phobic element for each patient, allowing a single device to be versatile for several patients with different phobias.

3.2 Integrating the therapist's workspace

One of our objectives is to take advantage of the space where the therapeutic session takes place and integrate the virtual triggering elements in an ecological way. This is to say that the virtual elements must be added in a coherent way to the space and existing furniture. For the therapist to take control, as above referred, there must be a common reference frame between his/her view and the patient's. This common frame enables the introduction of a virtual table on the therapist's view that corresponds to the physical table viewed by the patient. This allows the therapist to precisely choose where to place the virtual elements on this surface. As the 3D reconstruction of the environment in real-time can require a significant amount of computational power and may result in the appearance of disturbing artefacts, the choice fell on the superposition of models with some of the furniture elements. This way these elements will be represented in the therapist's view so that they can be used to precisely place those referred elements, while "invisible" in the patient's, as there is already the physical counterpart. This idea is accomplished with the use of fiducial markers (in this case, QR codes).

Figure 1: A user interacting with AR via Hololens 2.

3.3 Solution Architecture

The solution is divided into two parts: the patient's device and the therapist's device. First, the patient is able to see his phobic triggers through an augmented reality device. In our case, the Hololens 2 (figure 1) was picked since it is wire-free, lightweight, portable, and provides a high level of immersion. Besides these advantages it allows us to explore other aspects such as hand tracking capabilities, built-in voice commands, eye tracking, and spatial mapping.

On the other hand, the psychologist can control everything the patient sees in real-time. This includes managing the type, quantity, and movement of the phobic aspects visible to the user at every moment, allowing the creation of experiences tailored to each patient's needs. With the ability to choose the type of phobic trigger presented to the patient, the treatments of numerous phobias will also be possible with the same equipment.

Figure 2 shows two similar diagrams portraying the application's division into the two parts explained above. The patient has an immersive experience using the Hololens 2, viewing his phobic element in numerous ways. The psychologist controls everything the patient sees via a computer.

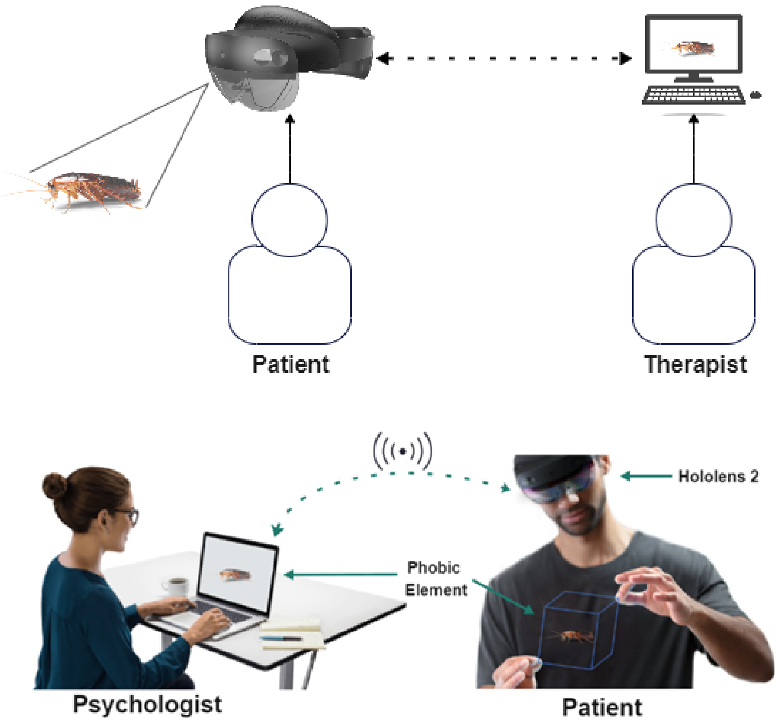

Figure 2: Illustration of the principles supported in the application.

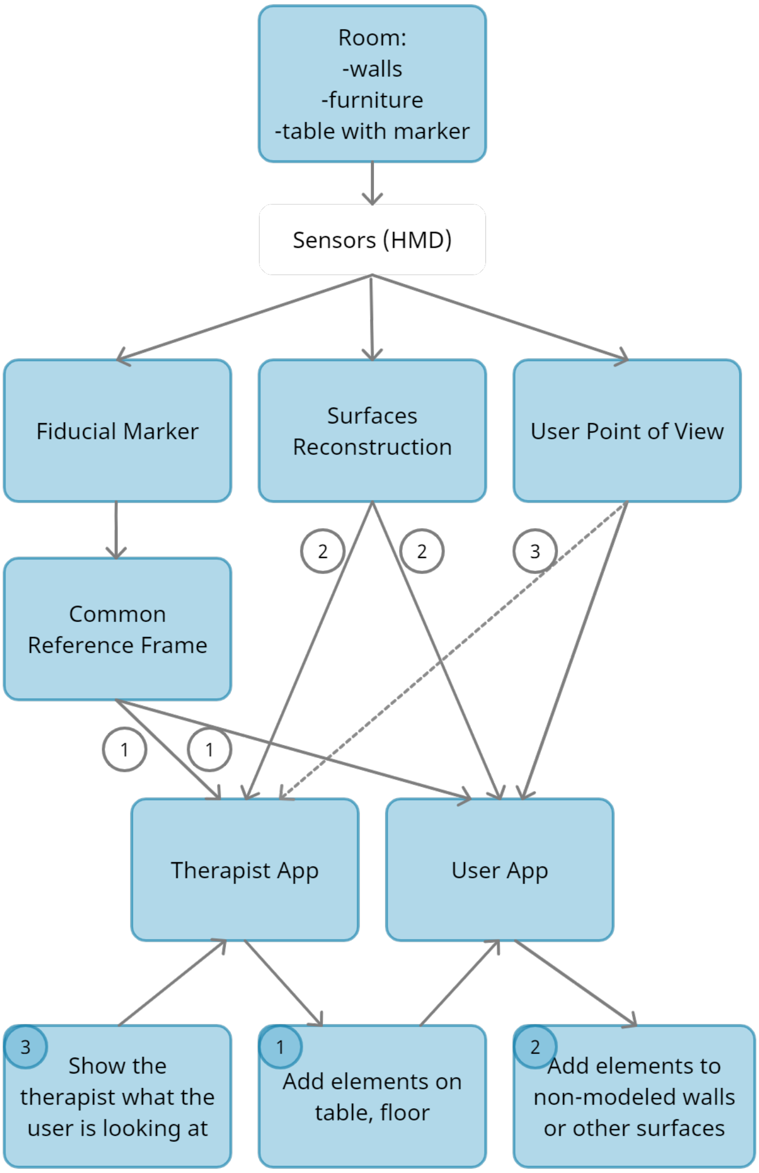

Figure 3: Flow diagram of the solution.

In figure 3 we can see the flow diagram of our solution. The HMD and its sensors can acquire a model of the room, the user's point of view, and perform surface reconstruction. The application includes fiducial marker detection and poses extraction to support the establishment of a common reference frame for both the HMD and the desktop applications. The reconstruction of the surfaces gives realism to the scene because the phobic elements can be put realistically on the objects in the room. The user's point of view is constantly monitored and used for the therapist to analyse in real time what is in the patient's field of view and if he/she tries to avoid looking at, or staring at some of the elements. All these features support this therapeutic system composed of the two already mentioned applications. This is made possible through the inclusion of appropriate communication support that will be described later. By putting all together, the virtual elements may be added, removed and controlled by the therapist, according to the interaction and intended stimulation for each patient at each specific moment of the exposure sessions.

4 Implementation and results

With the solution explained before in mind, we will now look at how we are carrying out the implementation of the project, explaining the modules used to put the ideas discussed into practice.

4.1 Development

As already mentioned, the target device chosen was Microsoft Hololens 2, which has an interesting set of sensors and capabilities, accessible through the Mixed Reality Toolkit (MRTK). The development for this device can be made directly using MRTK, or using a third party platform commonly used in game development such as Unreal or Unity. The choice fell on the latter, mostly driven by the existent experience within the research group. The development based on unity can take advantage of the vast number of assets available, both proprietary and opensource that support different types of functionalities like model imports, communications, or others. Examples of these assets are those provided by hardware vendors that support the development of applications for their specific platforms through some sort of hardware abstraction layer that hides the particular details. These assets, therefore, provide the necessary interfaces for the developer to extend them through his/her own assets and scripts.

These include scripts and components that help create mixed reality apps in an easy and quick way. This gives the developer freedom to use them or write his/her own assets and scripts, in the current case, build support for multiple device communication and the establishment of consistent points of view between an AR application and a 3D screen-based application, so the therapist can control the timing, location and intensity of stimuli generation.

4.2 Multiplatform communication

To allow communication between Hololens 2 and the therapist's computer, we are using Unity Multiplayer Networking Netcode, a Unity networking library that allows faster creation of networked applications. This uses UNet transport that relies on UDP transport layer to support the timed synchronization of the information between the devices.

Usability and flexibility concerns led to the implementation of a local network device discovery, to enable highly flexible usage of the setup, without relying on directory/game servers that would need permanent Internet connectivity. This discovery step is based on an initial broadcast of a discovery message initiated by the AR device (the client), expecting to receive a response from the therapist's computer (the server). After this initial phase, the peers are mutually identified, the connection is established supporting all the necessary communications used to place/move/destroy the phobic elements, parameter change, or other.

4.3 Establishment of a common reference frame

To enable the therapist to have a view of the patient's workspace/environment and interact with it while disposing of the elements as appropriate, a common reference frame should be established. It became also clear that some of the important physical elements should have a digital representation to make the above interaction possible. As the Hololens device establishes its own coordinate frame, it is necessary to infer one transformation per physical object that will have a virtual counterpart representation in order to support the proper display of the virtual representations in the therapist's view. Starting with a table, typically placed between the patient and the therapist, and where most of the interaction and exposure will be happening, a simple marker purposely placed on it can be enough support to the alignment of the digital model. This marker is detected on the images captured by one of the device's cameras and its pose estimation is consequently performed, enabling the necessary adjustment to the coherent placement of the virtual elements in front of the patient, through the therapist interface. In reality, although the marker detection and pose estimation requires that the patient "looks at it", once detected, and given that the Hololens has a quite stable pose estimation, the relative transformation is not expected to change even if the marker does not return into the camera field of view for a period of time.

4.4 Addition of the patient's hand reference point

Since the therapist cannot directly see the patient's view, it is difficult for the therapist to analyze the patient's interaction with the phobic elements. To mitigate this problem, it was necessary to detect the poses of the patient's hands and pass them to the therapist's application. Using this, the therapist observes in real time the patient's hands and their interaction with the phobic elements. In this way, the therapist can better analyze the patient's reactions and interactions with the phobic elements, adapting the exposure as needed.

4.5 Results and preliminary analysis

In figure 4 we can observe the result of the implementation explained above. In this image, we can understand the therapist's view of the exposure surface, and the patient's interactions, and therefore chose where to place the phobic elements. Both user and therapist can interact with the virtual objects, knowing that the position is updated for both perspectives. The fiducial marker is placed on an out of sight corner, and does not interfere with the experience. This is possible since the detection of the marker is only required in the beginning for the establishment of a common reference frame for the exposure table.

The design of the patient-side application was made to be simple and to minimize distractions for the patient, allowing the user to focus on the phobic elements, and on the therapist's instructions. The goal of the therapist-side application is to be easy to operate and intuitive. So, to add the animals just right-click on the place where you want them to appear. To remove

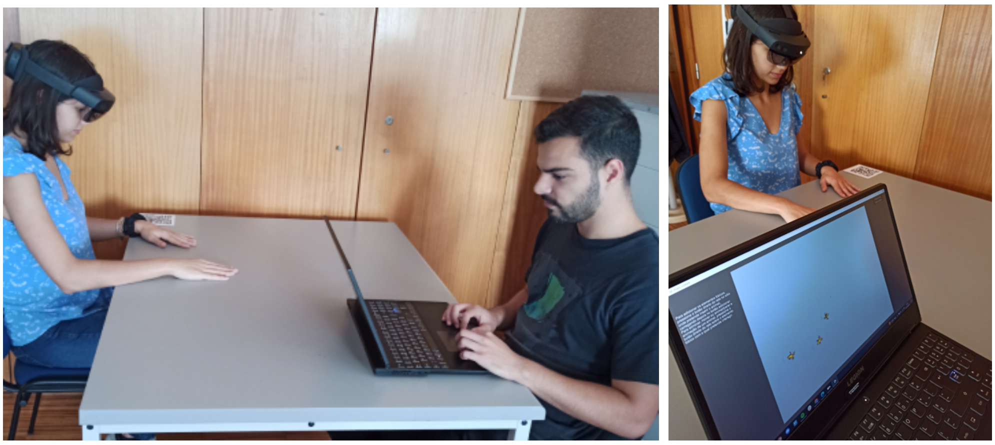

Figure 4: Representation of a possible therapy session with a patient using IS3TA on Hololens 2, whose exposure to the phobic elements is controlled by the therapist through the personal computer interface.

and move the animals, we perform the same operation on the animal we want to control at the moment. In figures 5 and 6 we can see the view from the patient with the two animals current present in the solution, frogs and spiders. The developed framework enables the easy addition of other types of phobic-triggering elements.

Although the system is still under development it was considered important to validate its current state. For that a preliminary study was made, where eleven non-phobic people were asked to try the AR application (i.e. the patient side of the solution) and answer a questionnaire about the immersion and quality of the AR environment. Although this test was performed with a reduced number of participants, it follows the ideas of Nielsen and Landauer [28], who stated that with only 5 users it is possible to find 85% of the problems in interactive systems. In fact, early identification of potential problems typically brings an enormous advantage, as corrections tend to be simpler to apply at this phase, rather than over a final product. The questions for this evaluation were chosen from the ”reality judgement and presence questionnaire” of Regenbrecht and Schubert [27] and are presented in the table below.

Figure 5: View from Hololens 2 with spiders as the phobic elements.

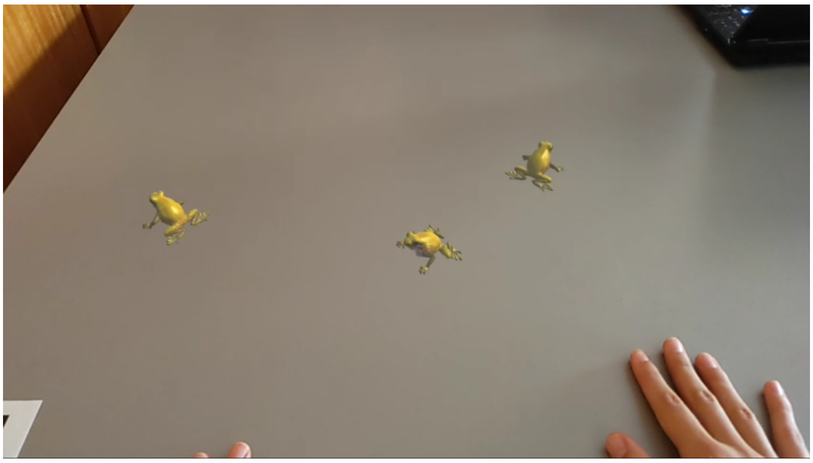

Figure 6: View from Hololens 2 with frogs as the phobic elements.

Q1 - Was watching the virtual objects just as natural as watching the real world?	1= No	5= Yes
Q2 - Did you have the impression that the virtual objects belonged to the real object, or did they seem separate from it?	1= Separate objects	5= Same object
Q3 - Did you have the impression that you could have touched and grasped the virtual objects?	1= No	5= Yes
Q4 - Did the virtual objects appear to be on a screen, or did you have the impression that they were located in space?	1= On screen	5= In space
Q5 - Did you have the impression of seeing the virtual objects as merely flat images or as three-dimensional objects?	1= Flat images	5= Three-dimensional objects
Q6 - Did you pay attention at all to the difference between real and virtual objects?	1= I Saw no difference	5= The difference was clear
Q7 - Did you have to make an effort to recognize the virtual objects as being three-dimensional?	1= Yes an effort was needed	5= No effort was needed

As explained before, the study was made with ten males and one female divided into two groups: the first group of six people who interacted with frogs and the second group of five people who interacted with spiders. The participants have ages ranging from 20 to 30 years old. The conditions of light were similar, and all people saw three animals on a table (first static and then animated to look alive). The Hololens 2 was also calibrated before every test to adjust the view of the models to each user. Figure 7 shows the average and standard deviation plots computed for the answers to the questionnaires separated into three series: total, frogs only, and spiders only.

From the analysis of the results obtained, we can conclude that the models are realistic and produce the desired effect of appearing to be part of the environment. From the observation of the values of questions Q1, Q2, Q3, and Q5, we can observe that the participants found the objects to have a

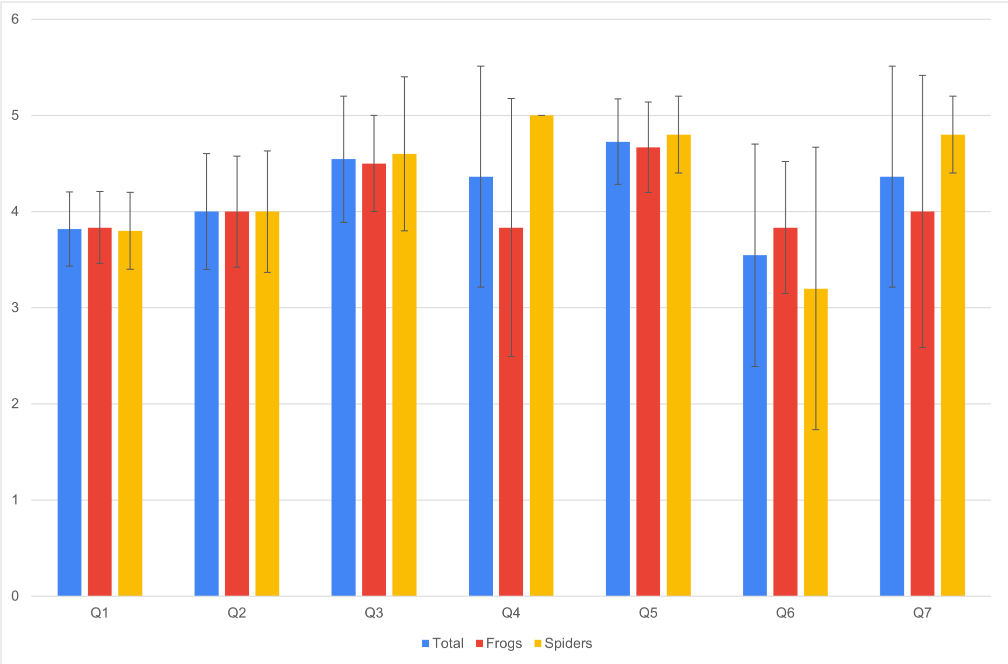

Figure 7: Average and standard deviation values for all the answers, frogs-only answers, and spiders-only answers.

three-dimensional and natural appearance, well integrated into the environment, and likely or appealing to be touched. There was no visible difference in the results between the two animal models used in the trial. Regarding question Q4, the volunteers who interacted with the frog models seemed to be somewhat aware of the screen used to create the holographic objects. When analyzing the values of question Q6, we understand that the difference between real objects and virtual objects was clearly perceived, however, the impact of this in the therapy has yet to be studied in a therapeutic context. At last, reviewing the results obtained in question Q7 we can conclude that, the virtual objects were recognized as being three-dimensional, especially for the spider models. In addition to these answers, it was also observed that as soon as the volunteers saw the models, most of them had an instinctive reaction to try to touch or smashed them, showing that the animals seem realistic.

One of the aspects reported, although already known to us, was the fact that the perception of the virtual elements depends on the ambient lighting. In reality, the holographic nature of the device is based on "adding light", and does not dispose of any means to block the light that comes from the objects. As a result, if the environment itself is too bright, the virtual elements tend to fade out.

Although this study does not address the therapeutic application context, it already brings some exciting results. The forthcoming steps will include the analysis made by a therapist and then, after incorporating possible recommendations, a clinical study will also be performed.

5 Conclusion

This paper proposes a solution to replace the exposure *in vivo*, using augmented reality. This method can suppress the problems associated with *in vivo* exposure, like the low acceptance and high dropout rates, the complications related to the presence of the phobic trigger and the therapist's inability to control it, as well as the impossibility of performing this therapy in the usual office environment.

Despite not being yet tested on people with phobias, this work shows promising results. In future work, we plan to improve it by adding additional features, like eye tracking and improved interaction.

References

[1] R.C. Kessler, W. T. Chiu, O. Demler, K. R. Merikangas, E. E. Walters,"Prevalence, Severity, and Comorbidity of 12-Month DSM-IV Disorders in the National Comorbidity Survey Replication", Archives of General Psychiatry, vol. 62, pp. 617-627, June 2005.

[2] U.S. Department of Health and Human Services,"Anxiety disorders", National Institute of Mental Health, Last accessed 4 January 2021, 2018, Avaliable online: https://www.nimh.nih.gov/health/topics/anxiety-disorders.

[3] S. Dattani, H. Ritchie, M. Roser,"Mental Health", Our World in Data, 2021.

[4] Y. Choy, A. J. Fyer, J. D. Lipsitz, "Treatment of specific phobia in adults", Clinical Psychology Review, vol. 27, pp. 266-286, 2007.

[5] N. Morina, H. Ijntema, K. Meyerbröker, P. M.G. Emmelkamp,"Can virtual reality exposure therapy gains be generalized to real-life? A meta-analysis of studies applying behavioral assessments", Behaviour Research and Therapy, vol. 74, pp. 18-24, 2015.

[6] C. Suso-Ribera, et al.,"Virtual Reality, Augmented Reality, and In Vivo Exposure Therapy: A Preliminary Comparison of Treatment Efficacy in Small Animal Phobia", Cyberpsychology, Behavior, and Social Networking, vol. 22, pp. 266-286, 2018.

[7] P.M.G. Emmelkamp, M. Krijn, A. M. Hulsbosch, S. de Vries, M. J. Schuemie, C. A. P. G. van der Mast,,"Virtual reality treatment versus exposure in vivo: a comparative evaluation in acrophobia", Behaviour Research and Therapy, vol. 40, pp. 509-516, 2002.

[8] T. D. Parsons, A. A. Rizzo,"Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis", Journal of Behavior Therapy and Experimental Psychiatry, vol. 39, pp. 250-261, 2008.

[9] I. C. Giglioli, F. Pallavicini, E. Pedroli, S. Serino, G. Riva, "Augmented Reality: A Brand New Challenge for the Assessment and Treatment of

Psychological Disorders", Computational and Mathematical Methods in Medicine, pp. 1-12, 2015

[10] N. A. J. De Witte, S. Scheveneels, R. Sels, G. Debard, D. Hermans, T. Van Daele, "Augmenting Exposure Therapy: Mobile Augmented Reality for Specific Phobia", Frontiers in Virtual Reality, vol. 1, pp. 8, 2020.

[11] M. Wrzesien, J. Burkhardt, M. Alcañiz, C. Botella, "How Technology Influences the Therapeutic Process: A Comparative Field Evaluation of Augmented Reality and In Vivo Exposure Therapy for Phobia of Small Animals", Springer Berlin Heidelberg, pp.523-540, 2011.

[12] C. Botella, W. A. Pérez-Ara, J. Bretón-López, S. Quero, A. García-Palacios, R. M. Baños, "In Vivo versus Augmented Reality Exposure in the Treatment of Small Animal Phobia: A Randomized Controlled Trial", PLOS ONE, vol. 11, pp. 1-22, February 2016.

[13] C. Botella, J. Fernández-Álvarez, V. Guillén, A. García-Palacios, R. Baños, "Recent Progress in Virtual Reality Exposure Therapy for Phobias: A Systematic Review", Current Psychiatry Reports, vol. 19, pp. 1-12, 2017.

[14] D. Boeldt, E. McMahon, M. McFaul, W. Greenleaf, "Using Virtual Reality Exposure Therapy to Enhance Treatment of Anxiety Disorders: Identifying Areas of Clinical Adoption and Potential Obstacles", Frontiers in Psychiatry, vol. 10, pp. 773, 2019.

[15] A. Garcia-Palacios, C. Botella, H. Hoffman, S. Fabregat, "Comparing Acceptance and Refusal Rates of Virtual Reality Exposure vs. In Vivo Exposure by Patients with Specific Phobias", CyberPsychology & Behavior, vol. 10, pp. 722-724, 2007.

[16] B. O. Olatunji, B. J. Deacon, J. S. Abramowitz,"The Cruelest Cure? Ethical Issues in the Implementation of Exposure-Based Treatments", Cognitive and Behavioral Practice, vol. 16, pp. 172-180, 2009.

[17] J. Szymanski, W. O'Donohue,"Fear of Spiders Questionnaire", Journal of Behavior Therapy and Experimental Psychiatry, vol. 26, pp. 31-34, 1995.

[18] R. M. Baños, C. Botella, M. Alcañiz, V. Liaño, B. Guerrero, B. Rey, “Immersion and Emotion: Their Impact on the Sense of Presence”, CyberPsychology & Behavior, vol. 7, pp. 734-741, 2006.

[19] A. Bangor, P. T. Kortum, J. T. Miller,“An Empirical Evaluation of the System Usability Scale”, International Journal of Human–Computer Interaction, vol. 24, pp. 574-594, 2008.

[20] B. O. Rothbaum, L. F. Hodges, R. Kooper, D. Opdyke, J. S. Williford, M. North, “Effectiveness of computer-generated (virtual reality) graded exposure in the treatment of acrophobia.” American Journal of Psychiatry, vol. 152, pp. 626-628, April 1995.

[21] S. Côté, S. Bouchard, “Documenting the efficacy of virtual reality exposure with psychophysiological and information processing measures.” Applied Psychophysiology and Biofeedback, vol. 30, pp.217-232, September 2005.

[22] C. T. Pitti, et al., “The combined use of virtual reality exposure in the treatment of agoraphobia.”, Actas Esp Psiquiatr, vol. 43, pp. 133-141, July-August 2015.

[23] M. C. Juan, M. Alcaniz, C. Monserrat, C. Botella, R. M. Baños and B. Guerrero, "Using augmented reality to treat phobias," IEEE Computer Graphics and Applications, vol. 25, pp. 31-37, November-December 2005.

[24] B. Patrão, P. Menezes, N. Gonçalves, “Augmented Shared Spaces: An Application for Exposure Psychotherapy.”, International Journal of Online and Biomedical Engineering (iJOE), vol. 16, pp. 43-50, 2020

[25] J. Bretón-López, S. Quero, C. Botella, A. García-Palacios, R. M. Baños, M. Alcañiz, “An Augmented Reality System Validation for the Treatment of Cockroach Phobia.” Cyberpsychology, Behavior, and Social Networking, vol. 13, pp. 705-710, December 2010.

[26] M. Wrzesien, M. Alcañiz, C. Botella, J. M. Burkhardt, J. Bretón-López, M. Ortega , D. B. Brotons, “The therapeutic lamp: treating small-animal phobias.” IEEE Comput Graph Appl, vol. 33, pp. 80-86, January-February 2013.

[27] H. Regenbrecht, T. Schubert, “Measuring Presence in Augmented Reality Environments: Design and a First Test of a Questionnaire”, arXiv, 2021.

[28] J. Nielsen, T. K. Landauer, “A mathematical model of the finding of usability problems. In Proceedings of the INTERACT ’93 and CHI ’93 Conference on Human Factors in Computing Systems (CHI ’93). Association for Computing Machinery”, pp. 206-213, 1993.