Singing for Anxiety: Exploring the Potentials of Singing in Immersive Virtual Reality to Help Children with Social Anxiety

Ali Adjorlu (adj@create.aau.dk) (0000-0001-5414-8631) and Stefania Serafin (sts@create.aau.dk) (0000-0001-6971-1132)

1Architecture Design and Media Technology, Aalborg University, A. C. Meyers Vænge 15, 2450 Copenhagen, Denmark

ABSTRACT

This paper presents an immersive Virtual Reality application designed for exposure therapy on children suffering from social anxiety. Two different interfaces for communication between the therapist and the client are designed and evaluated: One allowing a therapist to be immersed in the virtual environment with the client while wearing an HMD (3D user interface) and one where the therapist uses a traditional screen with only the client wearing an HMD (2D user interface). The therapist can use the interface to change the virtual scene in order to gradually expose the user to social singing situations in order to reduce anxiety.

A heuristic evaluation and a focus group interview were conducted on four clinicians, all experienced on working with children with social anxiety on a daily basis. According to the participants, the VR exposure therapy resulted in similar improvements as in-vivo exposure therapy [14]. As an example, a study conducted by Bouchard et al. on 59 participants showed that VR exposure therapy resulted in similar improvements in the Liebowitz social anxiety score as in vivo exposure therapy [15].

One of the most effective methods to treat anxiety is exposure therapy [4]. By gradually presenting the patient with stimuli that provoke their anxiety, the clinician seeks to remove the negative emotions through guidance and repetition. Exposure therapy for social anxiety typically involves exposing the patient to real-life social situations, such as giving a speech in front of an audience or attending a party. The psychologist will often be present with the patient, correcting their maladaptive behavior and beliefs and providing appropriate coping strategies.

Back in 1994, North et al. published one of the first known studies on using virtual reality to perform exposure therapy for individuals who suffer from acrophobia [9]. Since this pioneering study, a vast body of research has looked into using VR as a tool to treat a variety of anxieties, such as aerophobia [10], acrophobia [11], claustrophobia [12], and social anxieties [13]. A recent systematic review indicates that, in general, VR-based exposure therapy interventions for social anxiety have similar efficiency as in vivo exposure therapy [14]. As an example, a study conducted by Bouchard et al. on 59 participants showed that VR exposure therapy resulted in similar improvements in the Liebowitz social anxiety score as in vivo exposure therapy [15].

Singing, and especially singing in a choir, can have psychological benefits for individuals with chronic mental health issues, with research indicating a reduction in anxiety and an increase in positive emotions and relaxation, which translates to maintained and extended mental well-being [16, 17]. Music, as a universal language that transcends the barriers of verbal communication, provides a medium through which children can express emotions, build confidence, and foster social connections in a non-threatening environment. Research has demonstrated that engaging in musical activities, such as singing or playing...
an instrument, can significantly decrease levels of social anxiety in children by enhancing their self-esteem and social skills. Moreover, the structured yet flexible nature of music therapy allows for tailored interventions that meet the individual needs of children with social anxiety, encouraging gradual exposure to social situations in a controlled and supportive setting. However, research on the benefits of singing in virtual reality is in its infancy, with an acknowledged need for robust interdisciplinary studies [18].

Incorporating VR technology into music therapy presents a novel approach to treating social anxiety in children. VR-based musical interventions can simulate real-life social situations in a virtual environment, providing a unique opportunity for children to practice and enhance their social skills without the immediate pressures of actual social interactions.

Unfortunately, despite the benefits of VR as a tool for exposure therapy on individuals with SAD, no studies have been reported in recent systematic reviews on the topic with children as the primary target group for the intervention [14, 18]. In this paper, we present a VR intervention designed to perform exposure therapy for children with social anxiety by placing them in a virtual environment where they get to sing together with or in front of the Danish National Children’s Choir.

This paper presents a VR intervention designed to be a tool for therapists working with children suffering from social anxiety. However, this brief research report will focus on the needs and requirements of psychologists and therapists working with children and adolescents diagnosed suffering from SAD.

Brinkman et al. conducted a study focusing on the needs and requirements of the therapists rather than the client during a VR exposure therapy session [19]. This study resulted in a set of design guidelines when designing user interfaces (UI) for therapists to control a VR exposure therapy session. These guidelines focus on allowing the therapist to effortlessly control the VR scenarios experienced by the client and help them avoid activating unwanted events in the simulation. The design guidelines were for traditional screen-based UI. All of the studies described in the systematic review on VR for SAD conducted by Horigome et al. focus on the outcome of the users suffering from SAD [14]. However, several of these studies describe an interface designed for the therapist to control variables in the VR SAD exposure therapy interventions, such as the virtual avatars’ number of- and reactions. These interfaces utilize traditional interfaces such as a mouse and keyboard for the therapist to control the VR environment while only the clients are immersed in VR wearing a head-mounted display (HMD). Some scholars suggest that having the client wear an HMD can negatively affect the alliance between the client and the therapist due to the lack of eye contact [20]. This poses the question: what if the therapist is co-present in the virtual environment with the client while wearing an HMD and controlling an avatar, resulting in eye contact and gestural interaction in the virtual environment.

This paper presents a VR SAD exposure therapy intervention with a 3D UI interface designed for the therapist to use while controlling a 3D avatar and immersed via an HMD in the same virtual environment as the client. This intervention is then compared to a traditional UI controlled with a mouse and keyboard on a 2D screen to shed some initial light on the pros and cons of having therapists immersed in VR with their clients during exposure therapy.

2. METHODS

2.1 The VR intervention

The VR exposure therapy application consisted of eight 360° recordings of the Danish National Children’s Choir singing in one of their rehearsal halls at the Danish National TV studios (see figure 1).

A total of 25 children and one conductor participated in the recording sessions. In each video, the camera, the children, and the conductor’s locations were changed to capture scenes with varying levels of anxiety-inducing stimuli (see fig: 2).

- A scene where the camera is placed in the second row behind the rest of the choir members to create a socially safe VR experience. 

![Figure 1. Top: how the child sees the 3D avatar of the psychologist. Bottom: Singing session recorded from the second row.](image-url)
• A scene where the camera is placed in the first row creates a more anxiety-inducing experience than the previews scene.

• A scene where the camera is in front, next to the conductor, facing all of the children in the choir, creating the most anxiety-inducing experience.

• Three scenes were recorded where the nearest choir member’s distance from the camera are 40 cm, 60 cm, and 100 cm, to give the psychologist the option to create a different type of anxiety-inducing experience.

To explore whether the 360 recordings had the potential to be used as the main platform for the VR SAD exposure therapy intervention, we conducted a formative qualitative evaluation with two psychologists and one psychology student working in a clinic for children with anxiety. The psychology student is suffering from social anxiety. All three participants experienced the VR singing videos, followed by an unstructured group interview to get their views and suggestions for the intervention’s next iteration.

During the post-exposure interview, the psychology student with social anxiety stated that she felt a varying level of social anxiety having to sing in the back row compared to having to sing in front of all other choir members. She felt judged by the other choir members when placed facing them all and asked to sing. These statements from the psychology student do, to some extent, confirm that the videos might potentially be useful as a tool for exposure therapy on children diagnosed with social anxiety.

Both psychologists stated the importance of being able to change the amount of anxiety induced in the VR intervention gradually depending on the state of the client’s mind. They believed that the videos potentially could provide a varying degree of anxiety. They stated that since each child is different, they needed individualized treatment based on their specific needs, anxieties, and state of mind during the exposure therapy session. Changing scenes, pausing, fast-forwarding, or rewinding the 360 videos were among the psychologists’ requests to control the VR exposure therapy intervention.

Additionally, the psychologists mentioned the importance of trust and alliance between the child and the therapist during the exposure therapy session. The psychologists stated that the child should be able to point out anxiety-inducing elements in the virtual environment.

Based on this feedback, the next iteration of the application was developed. Three main functional requirements were identified:

• The ability to change the scenes in the virtual environment.

• The ability to pause, play, rewind, or fast-forward the 360 video.

• The psychologist should be able to communicate with the child using verbal and gestural interaction.

Additionally, in order to explore whether the alliance and communication between the psychologist and the client can benefit from having the psychologist immersed in the same virtual environment as the child, two versions of the interface were created (see figure 3):

1. **HMD Interface:** The psychologist is present in the virtual environment with the child, controlling a 3D avatar via HMD and VR controllers.

2. **Desktop interface:** The psychologist is present in the virtual environment with the child controlling a 3D avatar using the keyboard and mouse on a desktop computer (see figure 3).

The child will log in to the virtual environment using an Oculus HMD and touch controllers on both versions of the VR intervention. The application and the networking protocols are implemented in the game engine 3D Unity. The psychologist and the child can control their 3D avatar and 3D hand models, both of which are communicated over the network. This allows the psychologist and the child to see each other’s avatar body gestures in real-time. Only the psychologist can control the 360 video being played in the scene using either a 3D UI via the HMD interface or a traditional GUI in the desktop interface. These interfaces allow the psychologist to change the 360 video being played, pause, un-pause, fast-forward, and rewind. Additionally, the UI allows the psychologist to disable his avatar, giving the client the illusion of being alone in the virtual environment if needed.

### 2.2 Evaluation

Two psychologists and two psychology students, all working at a clinic for children suffering from social anxiety in Copenhagen, participated in an explorative evaluation of the application.

Participant A is an older male psychologist with many years of experience working with children and adolescents with mental disorders. He owns two clinics in Denmark, one in Copenhagen and one in Roskilde, both specialized in working with children with anxieties.

Participant B is a male psychologist working in the clinic owned by Participant A.

Participant C (Identifies as female) psychology student working at a clinic owned by Participant A. She is experienced in working with children with social anxiety.
Participant D is also a psychologist student and works as an assistant at the clinic owned by Participant A. Additionally, participant D suffers from a minor degree of social anxiety herself.

The study was conducted in the clinic owned by Participant A.

A within-group design with a predetermined shifted order for which UI version the participant should experience was conducted. Each participant tried both the desktop and the VR version of the application. A thorough introduction and walkthrough on how to use the interventions were given to each participant before they tried it themselves. Once they got to try the application themselves, one of the authors would log in to the app via the HMD on the second laptop and play the role of a child with social anxiety. The participants were given five minutes to try out the application with the author logged in as the child.

After five minutes, the author would provide the participants with a set of assignments. The assignments include changing to a specific scene, fast forward, pausing, unpause, and making the 3D avatar invisible or visible to the child. Following the task, the participants were asked a set of questions based on the revised set of heuristics proposed by Jakob Nielsen [21]. The chosen heuristics were visibility of system status, user control and freedom, error prevention and flexibility, and efficiency of use. Following the heuristic questions, a focus group interview was conducted with all four participants after trying both the desktop and the VR versions of the intervention.

2.3 Results

2.3.1 Heuristic Evaluation

The first heuristic was on flexibility and efficiency of use. In regards to the VR version of the intervention, all four participants reported that the system allowed them to control the exposure therapy session and interact with the child efficiently and flexibly. However, three out of the four participants (A, B, and D) reported that they had some issues fast-forwarding and rewinding the 360 videos when using the 3D UI.

For the desktop version of the application, three out of the four participants (A, B and C) believed that the system was efficient and flexible in supporting them in controlling the exposure therapy session while communicating with the child. Participant D stated that she had a hard time exploring the scene using the mouse and keyboard.

The second heuristic focused on error prevention. Regarding the 3D UI version, all four participants answered that they might make some errors during prolonged interaction with the system due to their inexperience with VR.

In regards to the error prevention of the desktop version of the intervention, three participants stated that they did not believe that they would make any mistakes while performing exposure therapy with the application.

The third heuristic was on the visibility of the system status. For the VR version of the application, three participants mentioned that they spent too much time trying to read the labels on the buttons, while one participant was interested in having an indication of how far they were into the movie despite there being a timeline on the 3D UI.

All four participants stated that they had a good overview of the system status via the user interface for the desktop version. One participant mentioned that this is due to him being more used to interacting with user interfaces using the keyboard and mouse.

The final heuristic was on user control and freedom. In regards to the 3D UI version, all participants mentioned that the system provided them the freedom to explore the scene effortlessly.

Regarding the desktop edition of the application, all participants stated that they were in full control of the exposure therapy session. Performing actions such as changing scenes was easy for all participants. Participant B added that even though he got control and freedom with this version, he missed out on experiencing the environment on the same level as the child (while wearing an HMD). Participant D stated that she struggled to explore the scene using the keyboard and mouse.

2.3.2 The focus group interview

Traditional coding [22] of the focus group interview resulted in four main categories: weaknesses and strengths of the desktop version of the application and weaknesses and strengths of the VR version of the application. Each of
In this paper, we set out to explore the benefits and disadvantages of allowing a psychologist to be immersed in the same virtual environment as the client when performing exposure therapy sessions. In order to do so, we created a SAD VR exposure therapy application with two interfaces for a therapist: One used while immersed in VR and one used via a traditional desktop computer.

During traditional in-vivo exposure therapy sessions, the psychologist is present with the client in the fear-inducing environment, observing the client’s behavior while providing appropriate guidance and feedback. Therefore, psychologists often analyze facial expression, body postures, and other parameters to interpret the emotions of their clients and then provide appropriate input based on these parameters. However, not all of these visual parameters are available when the client wears an HMD covering the face. What if the psychologist is also wearing an HMD and joins the client in the same virtual environment? In our implementation, when wearing the HMD, the psychologists and therapists could only see a 3D avatar that moved based on their clients’ head and hand movements. In general, according to the therapists participating in the short explorative study described in this paper, being immersed in VR enables the psychologist to create a greater sense of trust and collaboration with their client compared to when controlling the intervention via a desktop. Additionally, the VR version allows the client and the therapist to have eye contact in the virtual world, which is not possible when only the client is wearing an HMD. However, when the psychologist is wearing an HMD, some real-life information will be hidden, such as changes of skin color and sweat due to anxiety. In future iterations, biometric data can be provided to the therapist in real-time using devices such as the Empathica wristband [23] to increase the therapists’ ability to analyze the clients’ reactions while both are immersed in the intervention. This information can be intuitively presented to the therapist in the virtual environment.

The participants were, in general, more positive towards UI on the desktop version. He would have to move his left hand in front of his face in the VR version, covering the whole scene while he tried to press a button using the oculus touch controllers. Participants A and B both stated that it is much easier for them to see how the child feels since they could see the child’s body posture, skin color change, and sweat in real life when they were not wearing an HMD. Three out of the four negative comments regarding the desktop version of the application concerned the possibility of producing performance anxiety in the child if being the only one to wear an HMD. Participant D, who herself has some level of social anxiety, stated that when she was asked to wear the HMD, she felt nervous that she was the only one wearing the HMD and that others might negatively judge her. Finally, participants C and D stated that it was harder for her to navigate the scene via the desktop version than the application’s VR version. Moving her head to look around seemed more natural to her than having to use the mouse and keyboard to explore the scene.

3. DISCUSSION AND CONCLUSION

Figure 4. Results of the focus group interview. (Orange: Desktop UI. Blue: 3D UI)

These categories include several subcategories, which can be seen in figure 4.

There were a total of eight positive and seven negative comments on the VR version of the application. The most dominant topic among the positive comments was the power of VR to create equality between the psychologist and the child. Participant A stated that it could make the child feel safer if he knew that the psychologist is present in the virtual universe. Participant B added that it is crucial for a psychologist to put himself in the child’s shoes to create empathy. This is easier to achieve when experiencing the intervention wearing an HMD. Participant B stated that sometimes the child does not know why they are afraid of certain situations. By experiencing the scenario in the same way as the child using the HMD, it can become more manageable for the psychologists to discuss the anxiety-triggering events and offer an appropriate guidelines to the child on how to deal with them. Another positive topic discussed by the participants was trust. Participant D stated that if the child sees the psychologist wear the HMD and enter the virtual environment, it can increase the child’s trust towards the whole intervention.

Participant C stated that not being able to see parameters such as the child’s facial expressions and body language was one of the negative sides of both versions of the UI. Participant A states that these parameters are vital when doing in-vivo exposure therapy, giving VR exposure therapy a major handicap compared to it. Finally, during the focus group interview, participant B mentioned on three occasions that he felt some level of cybersickness while wearing the HMD. All four participants had negative comments regarding the user-friendliness of the 3D UI.

In regards to the desktop version of the application, a total of five positive comments were made during the focus group interview. Participant B stated that it was easier for him to keep an eye on what the child is doing while he tried to manipulate the exposure therapy session using the
the UI on the desktop version of the application compared to the UI for the VR version of the application. Previous experience using UI on desktop computers using the keyboard and mouse is the main reason for this preference. It can be argued that after some time of exposure and experience with interacting with user interfaces while wearing an HMD can increase the psychologists’ abilities in using a UI while wearing an HMD. However, there is still much work to be done when designing user interfaces to be used while wearing an HMD. As an example, due to the low resolution of the Oculus Rift HMD, it is not possible to read the labels on the buttons unless the UI is moved very close to cover most of one’s field of view. Newer HMD’s such as the Oculus Rift S or the HTC Vive Pro have a higher resolution that can fix this issue. However, future studies should investigate what methods to use to navigate and interact with graphical user interfaces while using VR controllers such as the oculus touch or the Vive Index. This study takes some essential initial steps towards understanding the needs of psychologists performing exposure therapy using VR instead of focusing on individuals suffering from social anxiety. Future research should further investigate the needs and requirements of psychologists and therapists when designing VR interventions for exposure therapy for children since they are one of the essential stakeholders.

Acknowledgments
We want to thank the Danish National Choir and the Danish National Radio for allowing us to record the sessions used in this paper. Additionally, we would like to thank Søren Benedikt Pedersen and his colleagues from “Cool Kids Copenhagen” for providing valuable feedback.

4. REFERENCES


