Creation of a virtual reality application for remote exposure therapy in social anxiety disorder

T. Langová[1], M. Jablonská[2], D. Jakeš[3], K. Adamcová[4], V. Trázníková[5], P. Stopková[6] & I. Fajnerová[7]

[1-7]Center for Virtual Reality Research in Mental Health and Neuroscience, National Institute of Mental Health, Topolová 748, Klecany, Czechia

[1]will.terka@gmail.com, [2]marketa.jablonska@nudz.cz, [3]david.jakes@nudz.cz, [4]katarina.adamcova@nudz.cz, [5]veronika.traznikova@nudz.cz, [6]pavla.stopkova@nudz.cz, [7]iveta.fajnerova@nudz.cz

[1-7]https://www.nimh.cz

DOI: 10.54337/aau.icdvrat26.19

Abstract

This paper describes the design and development of a self-guided VR application for remote exposure therapy targeting social anxiety disorder (SAD), referred to as FaceTheCrowd VR. It focuses on the design of interactive virtual environments and non-player characters enabling gradual exposure to socially challenging situations. The application is structured as a level-based experience in which users progress through scenarios of increasing social complexity within a simulated environment. The design of the exposure scenarios was based on previous studies identifying common anxiety-provoking social situations in individuals with SAD. The paper discusses the design considerations behind the virtual environments, character interactions, and technical implementation of the system, and outlines how these elements support the therapeutic goals of exposure therapy.

1. Introduction

Social anxiety disorder (SAD), also referred to as social phobia, is characterized by an intense fear of negative evaluation during social interactions and often leads to avoidance of everyday social situations (Schneier et al., 2015). SAD can be highly distressing, fundamentally affecting all areas of everyday life. The gold standard treatment for SAD includes pharmacotherapy and cognitive-behavioral therapy (CBT), with exposure therapy as a core component typically delivered in vivo (Canton et al., 2012).

Virtual reality exposure therapy (VRET) provides an alternative approach that has shown promise in replicating real-world scenarios while maintaining a safe and controlled environment (Horigome et al., 2020). Importantly, VR technologies enable the development of self-guided interventions that may improve accessibility of treatment for individuals with social anxiety. On the one hand, many individuals avoid face-to-face therapy due to personal barriers such as fear of social interaction, time constraints, or financial limitations. On the other hand, traditional exposure therapy is also limited by practical constraints, including the difficulty of recreating specific social situations, spatial limitations, a lack of available care providers, and long waiting times.

This paper presents the design and implementation of a self-guided VR application aimed at supporting remote exposure therapy for individuals with social anxiety.

2. Methods

This section describes the technical implementation and design of the proposed VR application. It outlines the hardware and software used to develop the system, the design of the exposure scenarios, sounds and non-player characters (NPCs), as well as the methods used for data collection.

1.1. Hardware and Software

The proposed solution is targeted at Meta Quest 2 and Meta Quest 3 headsets, as they are widely available and among the most used standalone VR headsets. The application is controlled using VR controllers, enabling the users to navigate the environments (via teleportation) and interact with objects.

The application is made in Unity game engine and uses XR Interaction Toolkit, which provides a high-level interaction framework for VR applications. Firebase online database is used to store user behavioural data outside the VR headset. The NPCs that appear inside the application are created in Character Creator 4.

1.2. Design of the Scenarios

The design of the virtual exposure scenarios was based on data from an online survey (N = 149) identifying common anxiety-provoking social situations experienced by individuals with social anxiety. The situations reported as most anxiety-provoking included public speaking in front of strangers, job interviews or debates, and situations involving public embarrassment. This was in parallel supported by the systematic review performed in our lab (Jakeš et al., under review). The scenarios (see Figure 1) are structured to allow gradual exposure, progressing from lower-intensity interactions to more socially demanding situations. Each scenario includes an introductory level that allows the user to explore the environment without other characters present and to practice the interactions used in the subsequent levels.

In the *Cafeteria* scenario, the user completes tasks such as entering and leaving the building, waiting for the server, ordering food, consuming it, and paying for the meal. The higher levels introduce more demanding situations, such as receiving an incomplete order or having to sit with another customer.

The *Interview* scenario consists of interactions with a receptionist, using an elevator, waiting in a waiting room, and subsequently answering interview questions for a specified duration of time. As the difficulty increases, additional characters appear in the building, interview room, and elevator. The user may encounter situations such as being told they arrived on an incorrect date, overhearing interviewers being rude to another applicant, or having to apologize for their phone ringing.

In the *Public Speaking* scenario, the user waits until they are invited to give a speech, approaches the podium, looks around the conference room, introduces themselves, and delivers a speech on a given (randomized) topic. In later levels, the user also answers questions from the audience. The difficulty primarily depends on the speech topics and the audience's reactions. In the second level, the user describes an image instead of being given a written topic, while in the third level the topics focus on the user's personal opinions and the audience members behave in a distracting or disengaged manner.

Figure 1 Screenshots of the scenarios. From left to right: Cafeteria, Interview, and Public Speaking.

1.3. Non-player Characters

The characters used in the application are created in Character Creator 4 (CC4) by Reallusion. As the application is intended to run on standalone VR headsets with limited computational resources, the characters have been converted to the Game Base format prior to export, which provides a more optimized representation suitable for real-time applications. Running the application with five non-optimized characters creates a significant reduction in frame rate, making the application difficult to use. Low frame rates in VR are known to increase the risk of simulator sickness, making performance optimization essential.

Each character is imported into the Unity game engine using the Unity Auto Setup plugin provided by Reallusion. After import, an Animator Controller component is assigned to the character to manage its animations. Most animations used in the application are sourced from ActorCore by Reallusion, as they are designed to be compatible with CC4 characters. Some animations were further modified directly in Unity to better fit the requirements of the application.

Based on the feedback from the validation study, the characters now also have facial animations (e.g. mouth movement, blinking, and facial expressions). These play an important role in social interaction and may influence the perceived realism of NPCs. This is particularly relevant for social anxiety, where some individuals are sensitive mainly to the presence of other people, while others respond more strongly to direct social evaluation and reciprocal interaction.

To improve scene loading performance, characters are instantiated after the scene is loaded using Unity's Addressables system. This approach also enables randomization of the characters present in the scenario. Randomizing the characters increases replayability, as users may encounter different characters during repeated sessions (e.g., sitting at a table with different NPCs in the Cafeteria scenario), which may be perceived differently by the user.

Characters with more complex behaviours are assigned a *BehaviorTreeRunner* component, which executes a behaviour tree (BT) stored in a Scriptable Object. Behaviour trees are commonly used for creating modular and reactive control systems for agents, enabling flexible switching between different tasks (Colledanchise et al., 2018).

The behaviour trees used in the application contain action nodes such as *ChangeAnimationBool, InvokeEvent, WaitForEvent, Wait, Speak,* and *WalkToPosition*, each with configurable parameters. For example, the Speak node specifies the audio clip that should be played while the node is active.

Character locomotion is controlled using Unity's *NavMeshAgent* component. Each scene therefore contains an object with a *NavMeshSurface* component with pre-baked navigation data.

1.4. Data Logging

During each scenario, two types of data are collected periodically (several times per second): data logging the user's behavior in the virtual environment, and events occurring within the application. This data is stored in a *.csv* file, which enables later automated data analysis.

The positions and rotations of the user's head and hands are tracked, together with the locations to which the user teleports within the environment, allowing reconstruction of the user's behavior during the scenario. To provide additional contextual information, the first object located in front of the user's head is also logged. As Meta Quest 2 and Meta Quest 3 headsets do not necessarily provide eye-tracking functionality, this serves as an approximation of what or whom the user is looking at. The volume of sound captured by the headset's microphone is also recorded, allowing approximate detection of whether the user is speaking.

As the scenarios contain quests for the user to complete, each quest completion event is logged, making it possible to connect user behavior with the current task objective. Another logged element is the name of the audio clip played at a given time interval, to see that an NPC is talking, some distracting sound is playing, etc.

Additional scenario-related information that is not bound to specific time intervals is stored in a *.json* file. This includes the participant ID (used to distinguish anonymous participants in studies), level name (indicating the scenario and its difficulty level), application build version, and NPC data (such as the names and/or genders of the generated NPCs). The file also contains psychological measures, including STAI-6 scores (State–Trait Anxiety Inventory values recorded before starting the scenario) and quest-related anxiety ratings measured using the Subjective Units of Distress Scale (SUDS) after completing the scenario.

All data is stored locally on the VR headset and, if a network connection is available, also uploaded to a Firebase server to reduce the risk of data loss (e.g., if the application is removed from the device).

1.5. Sounds

Sound plays an important role in increasing the sense of immersion within the application (Srithammee et al., 2025). Each scenario contains ambient environmental sounds appropriate to the given setting. For example, in the Cafeteria scenario the user can hear background sounds such as utensils clinking and people chatting. Higher difficulty levels incorporate additional distracting or potentially unsettling sounds, such as laughter, phones ringing, or people coughing, which are intended to increase the social pressure experienced by the user. The sound effects used in the application are obtained from online repositories providing audio under the Creative Commons Zero (CC0) license, such as Freesound.

Character dialogue is currently generated using the Narakeet text-to-speech (TTS) system as they provide Czech language TTS.

3. Results

Preliminary results from a validation study indicate that the VRET environment successfully induces anxiety in individuals with SAD compared to a non-phobic control group, supporting the ecological validity of virtual scenarios. In the validation of the Cafeteria scenario (experimental group: $n = 26$, control group: $n = 20$), differences in overall anxiety ratings (SUDS) between groups are statistically significant across all exposure levels (analyzed via non-parametric Mann-Whitney U test: level 0 – $U = 116$, $p = 0.001$, $rrb = 0.552$; level 1 – $U = 33$, $p < 0.001$, $rrb = 0.873$; level 2 – $U = 31$, $p < 0.001$, $rrb = 0.881$; level 3 – $U = 17.5$, $p < 0.001$, $rrb = 0.933$). Similarly, in the validation of the Interview scenario (experimental group: $n = 20$, control group: $n = 20$) statistically significant differences in SUDS ratings were observed at all exposure levels (Mann-Whitney U test: level 0 – $U = 118$, $p = 0.020$, $rrb = 0.410$; level 1 – $U = 33.5$, $p < 0.001$, $rrb = 0.848$; level 2 – $U = 43.5$, $p < 0.001$, $rrb = 0.782$; level 3 – $U = 38$, $p < 0.001$, $rrb = 0.800$). Both scenarios were rated by participants as easy to understand, providing an appropriately graded level of difficulty. A similar validation study of the Public speaking scenario is currently in progress. More detailed description of the VR application and validation results will be presented at the conference.

4. Conclusion

This paper presents the design and implementation of FaceTheCrowd VR, a self-guided VR application for exposure therapy targeting social anxiety disorder. The system provides a structured, level-based experience across three socially challenging scenarios of increasing complexity, designed to support the core therapeutic principles of gradual exposure. Preliminary results from a pilot validation study demonstrate that the application successfully differentiates anxiety responses between individuals with SAD and non-anxious controls, supporting the ecological validity of the designed virtual environments. These findings suggest that self-guided VRET represents a promising and accessible complement to traditional exposure-based interventions, particularly for individuals facing various barriers to in-person therapy.

The main contribution of this work is the design and implementation of a self-guided VR system for exposure therapy, together with a set of structured social scenarios and a data logging framework enabling behavioural analysis. Future work will focus on feasibility study for remote use, and evaluation of the effectiveness of repeated exposure sessions in social anxiety symptom reduction.

Acknowledgment

This work is supported by the project „Research of Excellence on Digital Technologies and Wellbeing CZ.02.01.01/00/22_008/0004583“ co-financed by the European Union.

References

Canton, J., Scott, K. M., & Glue, P. (2012). Optimal treatment of social phobia: systematic review and meta-analysis. *Neuropsychiatric disease and treatment*, 8, 203–215. http://doi.org/10.2147/NDT.S23317

Colledanchise, M., & Ögren, P. (2018). *Behavior trees in robotics and AI: An introduction*. CRC Press. http://doi.org/10.1201/9780429489105

Horigome, T., Kurokawa, S., Sawada, K., Kudo, S., Shiga, K., Mimura, M., & Kishimoto, T. (2020). Virtual reality exposure therapy for social anxiety disorder: a systematic review and meta-analysis. *Psychological Medicine*, 50(15), 2487–2497. http://doi.org/10.1017/S0033291720003785

Schneier, F., & Goldmark, J. (2015). Social anxiety disorder. In D. J. Stein & B. Vythilingum (Eds.), *Anxiety disorders and gender* (pp. 49–67). Springer International Publishing/Springer Nature. http://doi.org/10.1007/978-3-319-13060-6_3

Srithammee, N., Pannattee, P., Fukuchi, Y. & Nishiuchi, N. (2025). Investigating the Impact of Sound Design on User Experience in Virtual Reality. *17th International Conference on Knowledge and Smart Technology* (KST), Bangkok, Thailand, 2025. https://doi.org/10.1109/KST65016.2025.11003277