Exploring Methods of Rapid Exposure for Public Speaking Anxiety in Virtual Reality Exposure Therapy

Bradley Nolan

Supervisors: Dr. Mel Slater

Co-supervisor: Dr. Domna Banakou

Experimental Virtual Environments for Neuroscience and Technology

Departament de Psicologia Clinica i Psicobiologia

Universitat de Barcelona

Barcelona, Spain

Master Thesis in Brain and Cognition

Universitat Pompeu Fabra

July 2020

Table of Contents

Abstract

The use of exposure therapy in virtual reality as a treatment method for public speaking anxiety has been around for many years. Recent research has ventured into self-led therapy, as it has shown to be more ecologically valid than therapist-led therapy. The effectiveness of self-led therapy has not been completely solidified. To date, no study has been able to implement a self-led, rapid exposure method that would cause individuals to experience an environment with no audience, and perceive the presence of an audience, in a single session. The following report is divided into Experiment 1, the pilot study, and Experiment 2, the proposed follow-up study. Experiment 1 demonstrates a rapid exposure method. Experiment 1 was carried out on 9 individuals, in which 4 in the experiment condition experienced rapid exposure during a therapist session and 5 composed the control condition of simulating an in vivo therapist session. After the initial VR session, both conditions were to give a speech to a virtual audience as a follow-up session. The results of Experiment 1 showed no significant differences in the subjective measures between the two conditions. Factor analysis of a virtual reality questionnaire demonstrated high degrees of body ownership and agency in both conditions, along with a high degree of audience presence only in the follow-up session for both conditions. Experiment 2 proposes variations in the rapid exposure method, rooted in the concepts of transparency and proximity, and a control condition that experiences gradual exposure over multiple sessions. Experiment 2 aims to address limitations of Experiment 1 and contribute to this new line of research. More evidence is necessary before determining the efficacy of a self-led, single session virtual reality exposure scenario as a treatment method for public speaking anxiety.

Statement of Contribution

Experiment 1: Conceptualization, Methodology, Programming of the Virtual Environment, and Investigation were conducted by Domna Banakou and others from the EVENT Lab. I, Bradley Nolan, contributed to the Formal Analysis, Visualization, and Writing of the initial and final draft. Experiment 2: Conceptualization, Methodology, Programming of Virtual Environment were conducted by Domna Banakou and others from the EVENT Lab. I provided contributions towards the Methodology and Writing of the initial and final draft of the proposal. Both drafts were reviewed by Domna Banakou and Mel Slater.

Acknowledgements

I would like to acknowledge and thank Dr. Mel Slater for accepting me into the EVENT Laboratory to conduct my thesis. I would also like to express my gratitude to Dr. Domna Banakou for all the help and support I received throughout this entire process. Also, I would like to acknowledge the faculty of the Center for Brain and Cognition at the Universitat Pompeu Fabra for the accommodations and everything they have done for us in the Master's Program. Finally, my utmost appreciation goes to the entire master's class for their friendship and support. The environment created by the faculty, class, and city was more than an idealistic one for graduate school.

Introduction

Glossophobia. Fear of public speaking. Public speaking anxiety. Each of these terms has been used interchangeably to describe a subset of symptoms, or a specific type of social anxiety disorder (SAD)[1–4], in which individuals experience intense feelings of anxiety or fear when having to speak in front of a group of people. In this report, we have focused on the use of the term public speaking anxiety (PSA) to describe the anxiety experienced by participants during the experiment to be reported here. PSA, as the most common lifetime social fear[5], can potentially incapacitate individuals in their everyday lives, in which the fear and anxiety experienced become driving factors in creating new behaviors of avoidance towards public speaking[6,7]. One of the more common methods of treating PSA is cognitive behavioral therapy (CBT), which involves restructuring cognition through attempts to alter negative thinking and detrimental behavioral patterns[8]. Exposure therapy, a form of therapy that is frequently supplemented with cognitive behavioral therapy, attempts to desensitize individuals to stimuli that trigger anxious behavior, through gradual exposure to the stimuli over time. Recent evidence of its effectiveness has caused exposure therapy to become a widely used treatment for PSA, and even more so in immersive virtual reality[4].

The progress made in virtual reality (VR) and immersive virtual reality environments (IVREs) has facilitated the success of exposure therapy in a VR setting. Some of the previous work done has demonstrated the effectiveness of repeated, rapidly successive fear of public speaking scenarios. In recent work done by Takac and colleagues[1], the experiment consisted of three scenarios in which participants gave a speech in an IVRE to a new audience each time. The scenarios were capable of eliciting self-reported duress from a sample that did not have a clinical diagnosis for PSA. The initial speech had the highest level of distress reported, but after two more consecutive speeches, distress was effectively reduced, along with a slight, but insignificant, reduction in heart rate. Furthermore, research in the field has gone on to not only attempt to reduce distress, but to also reduce feelings of negative evaluation and speech anxiety as a treatment specifically. One such study provided evidence of self-reported improvements, where negative evaluation was reduced while positive self-assessment increased, even though no effect of distress reduction was observed over multiple sessions[3], which the insufficient sample used was reported to impact these results. Other work in the field has attempted other methods of therapy in VR, such as switching bodies with the therapist. A study done by Slater and colleagues demonstrated the application of this method[9]. As participants first spoke to a therapist about a personal problem, they then were given the opportunity to embody the therapist and listen and provide feedback to themselves, as if the participant was counseling themselves. The results depicted overall positive changes for the condition that switched bodies versus the condition that spoke with a scripted therapist.

More research has gone on to illustrate virtual reality exposure therapy (VRET) as a viable and ecologically valid treatment method for social anxiety and various phobias. A review by Coelho and colleagues compiled evidence that VRET can provide similar, and in some instances better, outcomes to that of in vivo exposure therapy for phobias such as the fear of heights[13]. Furthermore, one study by Landowska and colleagues explored prefrontal cortex (PFC) activity during VRET for the fear of heights[10]. The study used Functional Near-Infrared Spectroscopy (fNIRS), a non-invasive imaging technique that measures changes in the concentrations of oxygenated and deoxygenated blood in the brain through near-infrared light; the changes in these concentrations have been attributed to changes in neuronal activity. However, due to the non-invasive nature, fNIRS has poor spatial resolution in regard to specifying the exact location of brain activity; it is only able to record 4 cm deep. fNIRS makes up for this deficit with good temporal resolution, which permits an efficient method for tracking changes in brain activity over a specified period of time. The experiment demonstrated activity in the PFC was first diminished, depicting inhibition of the PFC and therefore more likely to experience fear and anxiety. However, activity later increased to normal with repeated sessions in an IVRE aimed at treating participants' fear of heights. Such evidence adds to the growing literature that supports the efficacy of VRET, observing neural changes that were correlated with subjective improvements. Because of the technological and scientific advancements made in the field, IVREs have become a valid and efficient medium in which exposure therapy is not limited by the physical constraints of natural environments and can be manipulated by clinical psychologists and other professionals to best suit their patients' needs. While these studies demonstrate certain scenarios to be effective applications of exposure and repetition, VR therapy also relies on effective immersion; certain concepts of immersive VR are integral for ensuring realistic responses.

For effective immersion into IVREs, a few specific factors must be observed. Primarily, they are the illusions of presence and body ownership. The illusion of presence is composed of two separate concepts: place illusion and plausibility illusion. Each of these refer to concepts that are possible within an IVRE and factor into the capability of a system to evoke natural responses from individuals. "Place illusion [is] the type of presence that refers to the sense of 'being there'... in spite of the sure knowledge that you are not there."[11] Place illusion is accomplished when an individual feels immersed and has strong feelings of being present in the virtual environment, almost to the point of forgetting about reality. However, place illusion can be lost. When it is lost, the individual experiences a break in the illusion, in which they recall the fact that they are interacting with a different reality through a head-mounted display (HMD). Most times, an event occurs that forces the individual to recall that they are still in the real world. Such events may be glitches that occur inside the system or may be sensory feedback from the environment outside of VR, such as interference of cords or contact with walls or objects. Additionally, place illusion may be domain specific, meaning the visual system may experience the illusion while the auditory system experiences breaks. While an auditory break affects place illusion as a whole, the visual system and illusion is left mostly unaffected. However, it is highly likely for place illusion to recover after the break, so long as the environment returns to its stable state or the individual reestablishes their bearings. The improving technology of HMDs also aids in providing the illusion of being there and limiting significant breaks in the system, which consists of everything from graphics, frame rate, latency, and screen resolution to body tracking and the physics of the IVRE.

Next, there is the plausibility illusion: "Plausibility is the illusion that, what is apparently happening is really happening (even though you know for sure that it is not)." To achieve the illusion of plausibility, a few aspects of the environment are taken together, these being the fact that events in the IVRE occur in relation to the individual, the IVRE responds and reacts to actions by the individual, and the individual believes the IVRE is credible. While the first two can be achieved through well-programmed environments, the credibility of the environment can vary between individuals. This is because it is dependent on individual expectations of the environment. If one individual believes what they are seeing to be an accurate representation of their reality, then they will feel the environment is authentic and therefore a higher level of plausibility. On the other hand, if an individual believes what they are seeing to be highly unlikely because they feel it is not representative of their reality, then the environment will feel flawed and produce a lower level of plausibility. However, this is not always viewed as detrimental. Some users of VR process the environment as not being credible compared to the real world but use that as a new understanding of the environment in which other actions, scenarios, or physics, are possible. This new understanding could update their illusion of plausibility, so long as the environment is consistent. Since these illusions are contingent on the individual subjective feeling of experience, or qualia, they cannot be objectively measured; they are instead inferred based on self-reports, the illusion of body ownership, and responses to the IVRE[12].

The concept of body ownership in VR originated with the Rubber Hand Illusion[13], an illusion that demonstrated the plasticity of the human brain in terms of proprioception, in which individuals have a sense of self-movement and body position and agency, where they feel responsible for bodily movements. The subjects of the Rubber Hand Illusion study, by Botvinik and Cohen, claimed to have a moderately high degree of ownership over the rubber hand through multisensory integration, in which touch sensation of their real, hidden hand was attributed to the simultaneous visual input of touch to the rubber hand. In addition, fMRI results have indicated that, when a threat was made to the rubber hand, brain activity in areas associated with anxiety, interoceptive awareness, multisensory integration, and motor function, all showed increases; the brain reacted with anxious and instinctual retraction responses similar to when participants' real hands were threatened[14]. The Rubber Hand Illusion was then further replicated in VR, demonstrating the capability of VR to elicit similar responses to the previous studies that used a physical arm. Sanchez-Vives and colleagues[15] and Slater and colleagues[16] demonstrated that synchronization of visuomotor stimulation, visuotactile stimulation, or both, are central factors for ensuring high degrees of body ownership in VR. The synchronization occurs when individuals visually observe their virtual body to be in unison with their movements (visuomotor) or when they visually observe touch on their virtual body to be in unison with what they actually feel with their corporeal body (visuotactile). Moreover, it was found that the virtual body is the strongest contributor to the plausibility illusion, when compared to other factors of an IVRE, including behavior of virtual avatars, physical coherence, and scenario coherence (credibility). Participants displayed an overwhelming preference for a full body-tracked avatar, with reports from the participants indicating the virtual body as being the most important aspect to improve when attempting to achieve the full illusion of plausibility[12]. The importance of these factors, of presence and body ownership, can be generalized to all types of IVREs, in that having high degrees of these help to limit the possibility of error through unrealistic or inconsistent behaviors.

The current study consists of reports of two experiments: Experiment 1 was a pilot study, and Experiment 2 is a proposed study that expands off of Experiment 1. The general goal of Experiment 1 was to explore the possibility of speeding up exposure therapy, that normally requires several sessions, through rapidly exposing individuals with PSA to an audience in a single session, rather than gradually over time. In the experiment, participants were set up to speak with a therapist, resembling Sigmund Freud[17], to discuss their PSA. During this first session, participants experienced a situation in which an audience emerged in the environment while simultaneously engaging in dialogue with the therapist. The audience emerged through a process of splitting, in which the therapist appeared to produce a copy of itself by having the copy come out of and split off from its body. Following the splitting process, the copies then underwent the process of morphing, in which the copies slowly changed and transformed their appearance from the therapist into a random human avatar. For the control condition, participants engaged in dialogue with the therapist and proceeded without any other effects occurring in the environment. As a follow-up session, participants in both conditions were then asked to return a week later and present a speech in front of a virtual audience. More details are presented in the Methods and Materials. The null hypothesis and hypotheses for Experiment 1 were as follows:

H_1: No differences will be observed between the experimental and control conditions for body ownership and agency measures.

H_2: Participants in the experimental condition will exhibit less anxiety and less fear of negative evaluation after the follow-up exposure compared to the control condition.

H_3: It is predicted that participants in the experimental condition will develop stronger feelings of perceived speaker competence for the follow-up session compared to the control condition.

For Experiment 2, the goal of the experiment is to test for an optimal method for the splitting process to simulate rapid exposure. The experiment will explore the possible influence of transparency and proxemics by comparing Near vs Far and Transparent vs Opaque. Therefore, four methods are proposed: Near+Opaque, Near+Transparent, Far+Opaque, Far+Transparent. When the splitting process begins, the therapist splits, at a moderately slow pace, and walks a few feet straight to one side nearby. This process is the same split as the one that occurred in Experiment 1. The first variation will alter the physical state of the new therapist during the splitting and will appear first as transparent, and then will become opaque as it progresses through the process. This may facilitate a less abrupt and shocking splitting process for the participant. The transparent morphing may cater more to the physics of a virtual environment and the overall acceptance of unnatural phenomena occurring. However, to date, no research has explored how transparency of virtual agents, especially through splitting, could affect individuals undergoing exposure therapy. The other variation proposed was to explore the effects of proxemics during the splitting process. Proxemics was a concept that originated in 1966 by Edward Hall, an American anthropologist[18]. Hall went on to describe that proxemics consists of four interpersonal distances: intimate (<1.5 ft/0.45 m), personal (4 ft/1.2 m), social (12 ft/3.6 m), and public (25 ft/7.6 m). This concept of proxemics has been previously demonstrated in VR, with participants exhibiting similar approaches to interpersonal distance for both human-controlled agents and computer-controlled avatars[19,20], and even inanimate objects when they are moving towards participants[21]. For this concept to be implemented in the proposed study, the newly copied therapist will instead turn and walk away from the original therapist and the participant, in a more diagonal route, with its destination residing in the public space, or near the border of the public and social space. With the copy appearing further away after splitting, it may alleviate some feelings of shock, anxiety, or distress due to personal space invasion during a therapy-like session. In addition, a control group will also be employed for Experiment 2, in which individuals will experience 6 sequential sessions of exposure therapy to simulate gradual in vivo exposure therapy methods. The first session will consist of only interaction and dialogue with the virtual therapist. Each sequential session following the initial will consist of an increasing number of virtual audience avatars, beginning with 3 and progressing to 15 by groups of 3. Some recent evidence has been provided regarding the effects of different virtual audience sizes on participants when public speaking in VR. It has been shown that a 3-avatar audience elicited a greater increase in HR compared to audiences with 6 and 15 avatars, respectively[22]. Therefore, rationale for beginning with 3 avatars and increasing to 15 is based on this previous work. Taken together, the current hypotheses and null hypothesis are provided:

H_1: No differences will be observed between experimental conditions or the control condition for measures of body ownership and agency.

H_2: The experimental conditions will elicit a greater change in anxiety than the control condition.

H_3: The Near proximity conditions will elicit a greater change in anxiety than the Far proximity conditions.

H_4: A difference in the change in anxiety will be observed between the Transparency and Opaque conditions.

As it is unknown what the participants in the study might prefer and why they might justify their preferences, no predictions have been made for the preference of experimental condition. However, it is still ideal for participants to report high degrees of body ownership and agency.

Experiment 1

Methods and Materials

Participants

10 participants were recruited for the pilot study through the University of Barcelona. These participants who agreed to take part in the experiment provided informed consent by reading and signing an ethical consent form. From here, participants were randomly assigned to the experimental (Freud Crowd) or control (Freud Only) condition. Demographic information was gathered from these participants, which included age, occupation, experience with computers, programming, VR, and video games. Except for age, the experience response variables were coded on a 1-7 scale, with 7 being the most. Occupation only reported undergraduate students, Masters students, or research assistants. The randomly assigned groups were comparable across variables. The values of central tendency, along with their central deviations, are recorded in Table A (see Appendix). Participants were screened using the PRCS to determine the level of confidence participants had before intervention. No exclusion criterium was implemented for PRCS scores. One participant was excluded from the final data analysis due to absence from session 2, providing a final sample size of 9. Information, consent forms, and questionnaires were provided in Spanish and Catalan. Translations to English were provided

for reporting and analysis. Participants received monetary compensation for their participation. The experiment was approved by Comissió Bioètica of Universitat de Barcelona.

Materials

The main equipment used was an HTC Vive (https://www.vive.com) head-mounted display and controllers. The head-mounted display (HMD) has a refresh rate of 90 Hz, and a resolution of 2160x1200 pixels (1080x1200 pixels per eye), its weight is 555 g. When wearing the HTC Vive participants can see the Virtual Reality scenario in 3D and in stereo with a 110º field of view as if they were using their head gaze normally to look around a scene. The device uses more than 70 sensors including a MEMS gyroscope, accelerometer and laser position sensors. The controllers are used to track participants' upper body movements and rotations in real time. The virtual environment was implemented using the Unity3D platform (https://unity.com/). The animation-enabled model of the virtual body including all other virtual characters in the scene (Freud and Crowd) were created with Mixamo Fuse (https://www.mixamo.com/fuse) and customized appropriately for the purposes of the study using Mudbox 2019 and 3ds Max 2019 from Autodesk (https://www.autodesk.com) academic versions.

Measures

The following measures used in Experiment 1 have been chosen based on prior evidence of reliability and validity when measuring aspects of PSA, particularly fear of negative evaluation from other individuals, anxiety occurring before or during public speaking, and the self-perceived competence of one's own ability to give a quality speech[23]. The study incorporated an implicit measure for PSA and subjective questionnaires for both PSA and VR.

Personal Report of Confidence as Speaker (PRCS)

The PRCS is a 30-item self-report measure[24] that assesses speech anxiety with a high degree of reliability and internal consistency[25] Sample questions included, "While preparing a speech I am in a constant state of anxiety," and "My thoughts become confused and jumbled when I speak before an audience." Responses were scored based on the extent of agreement on a scale from 1 to 6, with higher scores indicating a greater degree of anxiety. In the experiment, the

PRCS was only used once as a screening measure to determine the degree of confidence participants possessed before beginning the experiment.

Brief Fear of Negative Evaluation (BFNE)

The BFNE is a 12-item scale measuring anxiety concerning social evaluation[26]. Items are rated on a 5-point Likert scale ranging from 1 (not at all characteristic of me) to 5 (extremely characteristic of me), with a total score of 60 used for analysis. Higher scores on the BFNE indicate more negative evaluation and anxiety. In the experiment, the BFNE was used as an overall pre-test/post-test measure, only being conducted prior to the first session and following the second session. Because of this, a single variable for the change in BFNE score (*dBFNE*) was utilized for analysis, using the pre-BFNE mean score and subtracting it from the post-BFNE mean score (*dBFNE* = postBFNE-preBFNE). Therefore, a negative *dBFNE* is seen as improvement and a positive dBFNE as an exacerbation of anxiety.

Implicit Association Test (IAT)

The IAT is a test that measures the strength of associations between concepts by comparing response times on two combined discrimination tasks[27]. Concepts appear in a 2-choice task and the attribute in a second task. For this study, the IAT was used, which was adapted from the original and focused on testing for implicit self-esteem[28]. It is based on the choice to categorize concepts based on whether they apply to oneself or to others.

The IAT was assessed before intervention to acquire a baseline measure. Then, participants completed the IAT after sessions 1 and 2. To calculate the change in IAT over the four measurement points, baseline IAT score (preIAT) were subtracted from the post-IAT scores (*dIAT* = postIAT-preIAT; *dIAT1*, *dIAT2*). The IAT is scored on a scale from -1 to 1, utilizing the d-scores obtained from the test. Due to the fact that lower IAT scores are associated with “Others-Liked” and “Me-Rejected” and higher IAT scores are associated with “Me-Liked” and “Others-Rejected”, negative *dIAT* scores would indicate more anxiety and fear of negative evaluation, while higher scores would indicate less anxiety and fear.

State Trait Anxiety Inventory (STAI)

The STAI – State Version is a 20-item scale designed to measure present anxiety[29]. However, a 6-item inventory was supported by high internal consistency and correlation with the 20-item version[30]. This shortened version was used for the present study. Items for this experiment were rated on a 6-point Likert scale ranging from 0 (not at all) to 5 (very much so). Higher scores on the STAI indicate higher levels of state anxiety. Because the STAI was reported at four separate time points (pre/post sessions 1 and 2), the change in mean score was calculated for each session, using the pre-STAI mean scores and subtracting them from the post-STAI scores (*dSTAI* = postSTAI-preSTAI). Therefore, negative *dSTAI* scores are seen as decreases in state anxiety while positive *dSTAI* scores indicate increases.

State Perceived Index of Competence (SPIC)

The SPIC was used as a measure for perceived competence to complete a task[31] and has also been used in VR studies aimed at treating PSA[32] In the present study, it was to measure participants' self-perceived competence to give a public speech. In the experiment, the SPIC was used as a pre-test/post-test measure, only being conducted immediately following the first session and preceding the second session. A single variable for the change in SPIC score (*dSPIC*) was utilized for analysis, using the pre-SPIC mean score and subtracting it from the post-SPIC mean score (*dSPIC* = postSPIC-preSPIC). Higher scores on the SPIC indicate less perceived competence to complete the task. Therefore, a negative *dSPIC* is seen as an increase in perceived competence, for the week between session 1 and 2, while a positive *dSPIC* would be seen as a decrease in perceived competence over that time.

Virtual Reality Questionnaire (VRQ)

The Virtual Reality Questionnaire (VRQ) is 8-item Likert-scale questionnaire that incorporates questions used previously in VR experiments[9,17,33,34] The VRQ includes statements in order to test for body ownership, agency, and characteristics of the audience. Scored from 1 to 7 in terms of agreement, with 1 being "Not at all", 4 being "Neutral, and 7 being "Completely [agree]". The questions are included in Table G with variable names corresponding to responses after session 1 and after session 2, respectively.

Experimental Design and Procedure

Participants who were recruited and brought to the laboratory were provided with basic information about the experiment. After, participants were given a written informed consent form and were reminded that they were free to withdraw from the experiment at any time. Then, participants were asked to fill out the questionnaire for demographic information, the PRCS-12, the BFNE, and the IAT. The participants then registered for a pre-arranged timeslot for session 1 of the experiment. Upon arrival, participants were required to complete the STAI and the SPIC. Before the first VR session, participants were randomly assigned to either the experimental condition or the control condition. Once assigned, participants were immersed in the IVRE through the HMD. In both groups, participants were embodied in a same-gender virtual body that moved synchronously with their real movements.

Once in the environment, the participants found themselves talking to a virtual therapist (Sigmund Freud). The virtual therapist was programmed to talk with the participant about their fear of public speaking, and it was identical for both groups (for example: “Do you have to give public speeches often?”, “When was the last time you gave a public speech?”, “What kind of thoughts do you have about yourself when you think about giving a public speech?”, etc.). Participants were asked to participate in some public speaking exercises, including counting to 50, singing, describing a hobby or the synopsis of a book, movie, or television show. During the exercises, the therapist began to divide and copy himself, repeatedly, until eventually a group of copies has appeared in the environment. Those copies of the therapist also morphed and changed into different, random human characters. While morphing was occurring to the new copies, the new human characters eventually take seats in the chairs behind the therapist (Figure 1).

Figure 1: Experimental conditions: (**A**) the experimental/morphing group and (**B**) the control group.

In the control group the participants went through the same process of talking with the therapist, but there was no emergence of the audience. After the end of the session, participants immediately provided their responses for the STAI, IAT, SPIC, and the VRQ. Approximately one week later, the participants returned for the virtual audience follow-up experience. Participants were allotted 5 minutes to plan a structured talk based on responses they provided during the first session. This follow-up session consisted of a scenario in which participants will found themselves in front of a virtual crowd, and they were asked to give their planned talk. Once complete, participants exited the IVRE and recorded their responses for the BFNE, IAT, STAI, SPIC, and VRQ. The participants were then be debriefed on the study and compensated for their participation.

Data Analysis

All data processing and analyses were conducted in Python, accompanied by various libraries[35–39]. Preliminary analyses were conducted on the pilot data to determine if violations of the assumptions of normality and homoscedasticity were present. Only the raw variable scores from the VRQ violated Levene's assumption of normality. For the PRCS, *dBFNE*, and *dSPIC*, independent samples T-tests were conducted on the mean scores of the two experimental condition. For *dIAT* and *dSTAI*, 2 x 2 mixed-factorial ANOVAs were conducted, with session (1 & 2) as the within-factor and condition (Freud Crowd & Freud Only) as the between-factor.

Exploratory analyses were conducted on the VRQ to examine its structure and attempt to reduce its dimensions for better explanation of the responses. Results did not violate assumptions of normality or homoscedasticity.

Results

The following results were grouped based on main concepts for PSA: Negative evaluation, Anxiety, and Speaker Competence. Negative evaluation was based on the scores from *dBFNE* and *dIAT*; Anxiety was based on the scores from *dSTAI*; Speaker Competence was based on the scores for *dSPIC*. Raw scores for all measures were provided in Figure A (see Appendix). As the PRCS was only conducted as a screening measure, it was not included in the analysis for Speaker Competence. However, as all participants scored below half on the PRCS, all participants self-reported moderate-to-high levels of confidence for speaking. Mean values from

Figure B indicated that the experimental condition and control condition each had relatively equal scores, demonstrating no significant difference in the T-test before intervention (T = 0.406, p = 0.70, Cohen's d = 0.268).

Negative Evaluation

A slight trend was observed in Figure 2A, with a decrease in mean BFNE for the experimental condition (Freud Crowd: *M* = -0.80, SE = 2.08) and an increase for the control condition (Freud Only: *M* = 2.00, SE = 3.63), indicating slightly more fear of negative evaluation after session 2. However, no difference was observed from the T-test for *dBFNE* (T = -0.669, p = 0.27, Cohen's d = 0.473). Similarly, the ANOVA for *dIAT* did not produce a significant interaction ($F(1,7) = 0.291$, p = 0.61, $\eta^2 = 0.016$) or significant main effects (ps > 0.29). In Figure 2B, a larger decrease for *dIAT2* can be seen for the Freud Only condition compared to the Freud Crowd condition. However, this difference is not significant (T = 1.273, p = 0.24, Cohen's d = 0.813).

Anxiety

The ANOVA for *dSTAI* did not produce a significant interaction ($F(1,7) = 0.311$, p = 0.59, $\eta^2 = 0.023$) nor significant main effects (ps > 0.68). Per Figure 2C, it can be seen the experimental condition presented a trend in which state anxiety slightly increased throughout the entire experiment, while the control condition saw a slight decrease.

Speaking Competence

In Figure 2D, for both conditions, an increase in mean SPIC score was observed (Freud Crowd: *M* = 6.00, SE = 3.11; Freud Only: *M* = 1.00, SE = 1.73), with the experimental condition demonstrating a greater increase in *dSPIC* and therefore overall less perceived competence than the control condition. and *dSPIC* (T = 1.403, p = 0.20, Cohen's d = 0.872) between the two experimental conditions.

Virtual Reality

It can be seen in Figure 3 and Table 1 that the majority of participants in both conditions reported above neutral scores, with the exception of *Presence* and *Audience* during session 1 for the Freud Only condition. This exemplifies the high degree of body ownership and presence necessary to ideally achieve reliable results in VR experiments. Due to the violations of normality and the fact that the VRQ are ordinal variables, non-parametric one-way ANOVAs were conducted for session 1 and 2 separately, using condition as the between-factor. The Kruskal-Wallis H-test demonstrated non-significant effects for both conditions (Session 1: H = 3.003, p = 0.08; Session 2: H = 1.214, p = 0.27). None of the post-hoc pairwise Mann-Whitney U-tests proved to be significant (all ps > 0.15; Table H). All effect sizes were converted and supplied in Table 2, separated by session. Nearly half of the variables (7/16) carry medium effect sizes. For the exploratory analysis, it was observed in Figure C that strong, significant correlations were present amongst the VRQ variables, namely between *VRBody* and *Agency* (r_s = .812) and also between *Presence* and *Audience* (r_s = 0.918). These results prompted exploration through a principle component analysis on the 8 variables of Table 1 (n = 144) in an attempt to reduce the dimensions and further examine the correlated variables. The initial analysis indicated the first principle component explained 48.2% of the variance and the second explained 22.4%, for a total of 70.6%.

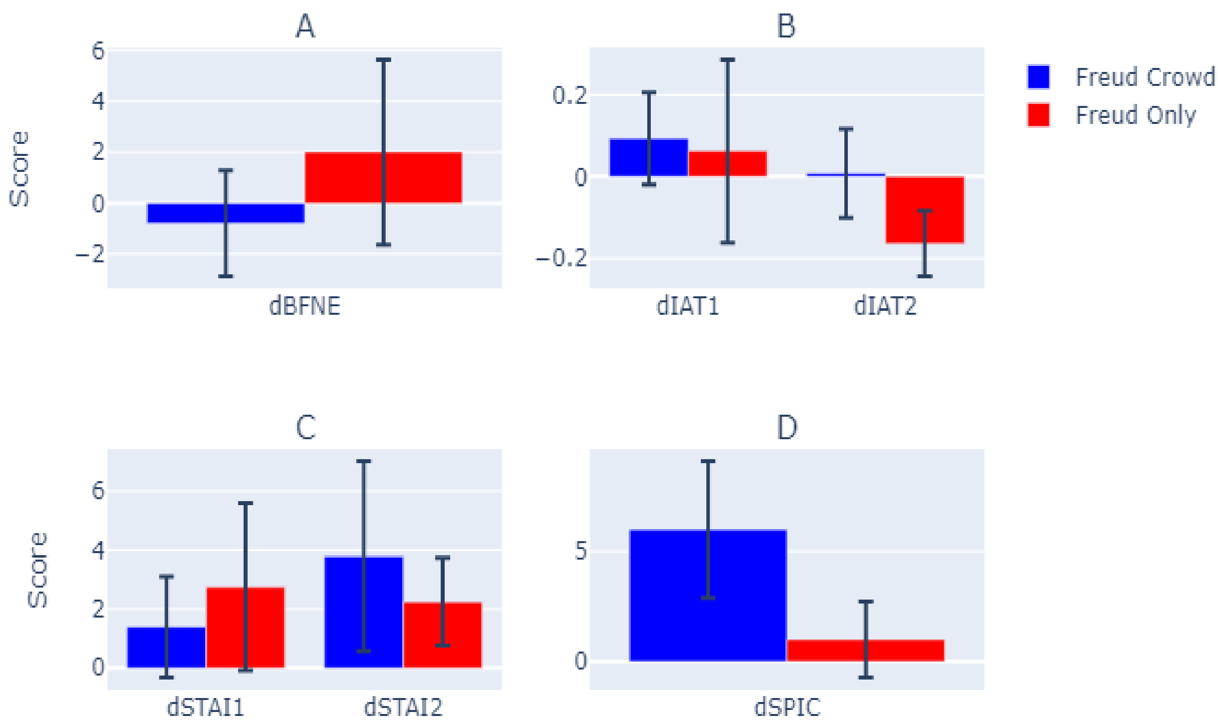

Figure 2. The bar charts for the change in scores for the two experimental conditions. (**A**) Bar chart showing the means and standard errors of *dBFNE*. (**B**) Bar chart showing the means and standard errors for *dIAT* for sessions 1 and 2. (**C**) Bar chart showing the means and standard errors for *dSTAI* for sessions 1 and 2. (**D**) Bar chart showing the means and standard errors for *dSPIC*.

From there, a factor analysis was conducted on the two factors utilizing the Maximum Likelihoods method and PROMAX oblique rotation. 27.0% of the variance was explained by Factor 1 and 54.0% of the variance was explained by both Factor 1 and Factor 2. Table I suggests that Factor 1 is highly dependent on *Presence* and *Audience* and Factor 2 on *VRBody* and *Agency*. This is confirmed in Table J as the factor variables were strongly correlated with the originals (ps < 0.001), therefore Factor 1 was termed Y_{aud} and Factor 2 as Y_{body}. For Y_{aud}, a significant difference was observed between sessions 1 and 2 ($F(1, 7) = 11.338$, $p < 0.01$, $\eta^2 = 0.586$). As the scores for Y_{aud} were greater for session 2 compared to session 1, this validated the fact that session 2 produced an audience for participants in both conditions. For Y_{body}, no difference was observed between sessions 1 and 2 ($F(1, 7) = 0.359$, $p > 0.56$, $\eta^2 = 0.043$). This indicated that feelings of body ownership and agency stayed consistent across sessions, regardless of condition.

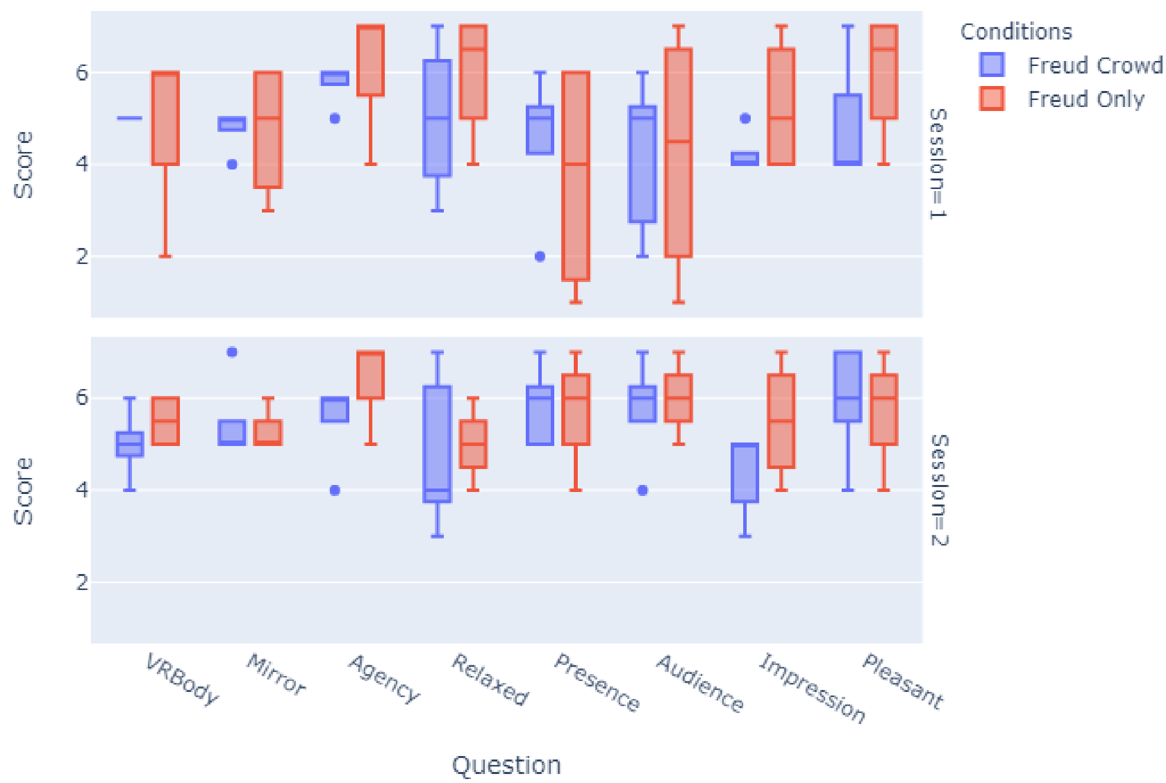

Figure 3. Box plots for the VRQ, corresponding to the variables in Table G. The top graph corresponds to the scores after session 1 and the bottom graph corresponds to the scores after session 2. The horizontal lines inside the box are the medians, the boxes are the interquartile ranges (IQR), and the whiskers range between [lower quartile – 1.5*IQR] and [upper quartile + 1.5*IQR]. Values outside the IQR are displayed as individual points. The possible range for scores is from 1 to 7, with 4 as neutral.

	VRBody	*Mirror*	*Agency*	*Relaxed*	*Presence*	*Audience*	*Impression*	*Pleasant*
Session 1 Freud Crowd	5 (0)	5 (0)	6 (0)	5 (2)	5 (0)	5 (2)	4 (0)	4 (1)
Session 1 Freud Only	6 (1)	5 (2.25)	7 (0.75)	6.5 (1.50)	4 (4.25)	4.5 (3.75)	5 (2.25)	6.5 (1.5)
Session 2 Freud Crowd	5 (0)	5 (0)	6 (0)	4 (2)	6 (1)	6 (0)	5 (1)	6 (1)
Session 2 Freud Only	5.5 (1)	5 (0.25)	7 (0.50)	5 (0.50)	6 (0.75)	6 (0.50)	5.5 (1.50)	6 (0.75)

Table 1. Median (IQR) for the variables for the VRQ, corresponding to Figure 3. FO = Freud Only, FC = Freud Crowd.

Variable	**Session 1**	**Session 2**
VRBody	0.408	-0.357
Mirror	0.024	-0.099
Agency	-0.212	-0.431
Relaxed	-0.314	-0.074
Presence	0.201	-0.024
Audience	-0.011	-0.101
Impression	-0.451	-0.453
Pleasant	-0.406	0.100

Table 2. Effect size statistics converted from Cohen's d scores from Table H. The effect size describes the non-parametric relationship between the median scores of the Freud Only and Freud Crowd conditions. Negative values indicate that the Freud Only condition is greater than the Freud Crowd condition.

Experiment 2

Methods and Materials

Participants

Participants with ages ranging between 18 and 60 years are eligible to take part in the present study. The sample size desired for this study will be 25 participants per condition (100 participants in total). This is based on a power analysis with power set to 80% and a predicted η^2 of 0.1[37]. All participants must manifest minimum a minimum level of anxiety from speaking in public and/or social anxiety, which might generate moderate to high levels of distress when exposed. For this reason, candidates will undergo a screening process to identify those eligible to participate. The Brief Personal Report of Confidence as Speaker (PRCS-12)[25] and the Liebowitz Social Anxiety Scale (LSAS)[40] will be used for the screening. The PRCS-12 has been validated with the Spanish population and it will be used to guarantee a minimum level of PSA in order for candidates to be able to participate to the experiment (with a cut-off point of 32, indicating moderate to high PSA). The LSAS scale will be used to assure that only participants with social anxiety related to PSA will take part in the study, and to avoid recruiting participants with other social phobias that could lead to confounding results (with a cut-off point of 30, indicating moderate levels of social anxiety)[41]. Participants will be compensated upon completion of the experiment.

Materials

The same apparatus described for Experiment 1 will be used for Experiment 2. There may also be the possibility of implementing use of the Oculus Quest for Experiment 2, as the Quest is currently becoming the most widely used HMDs on the market[42], due to it being a powerful, standalone HMD with hand-tracking.

Measures

Measures that were used in Experiment 1 (IAT, SPIC, and VRQ) will follow the same structure and scoring. Measures that are different from Experiment 1 are provided as follows:

Brief Personal Report of Confidence as Speaker (PRCS-12)

The PRCS-12 is a 12-item self-report measure that assesses speech anxiety. The current version of this scale[25] was derived from the 30-item scale[24] to include the 12 most predictive items. The PRCS-12 is measured in the same manner as the 30-item scale, with higher scores indicating more anxiety related to speaking. This measure will be used in place of the PRCS used in Experiment 1 as a screening measure to reduce the burden of this questionnaire, as the shortened version has been shown to be highly consistent and reliable when measuring for speech anxiety[25].

Liebowitz Social Anxiety Scale (LSAS)

The 24-item self-report version of the LSAS will be used to assess social anxiety[40]. The LSAS assesses fear and avoidance of various social situations, including performance situations such as public speaking. Each social situation is rated on 7-point scales for degree of fear (0 "none" to 6 "severe") and degree of avoidance (0 "never" to 6 "usually"). The incorporation of the LSAS is to get a better sense of subjective reports of avoidance in relation to social anxiety as a baseline measure.

Subjective Units of Distress Scale (SUDS)

SUDS, a scale developed in 1966 by Wolpe and Lazarus[43], has been previously used within exposure therapy[44] to gather real time anxiety ratings, in particular in VRET[45,46]. The scale is rated on a scale from 0 to 10, with 10 indicating strong feelings of distress and anxiety. The aim for Experiment 2 is to use the SUDS to retrieve multiple measures of anxiety throughout the experiment in order to better determine the causes for the change in anxiety. Participants will also self-report why they chose the SUDS score at each timepoint. Eight total measurement points are to be used during the first session, specifically targeting the public speaking exercises and the splitting and morphing that occurs in the experimental conditions.

Post-VR Interview

The post-VR interview consists of 10 open-ended questions that will be paired with the VRQ to further extrapolate information from participants regarding their subjective feelings for the experimental conditions. As some questions will result in count responses, they will have the potential to be converted, coded, and grouped into variables that can be included in the analysis of the VRQ. Questions are provided in the Appendix.

Behavioral Measures

In addition to the subjective measures, external experts will provide analyses based on video footage of the entirety of the experimental sessions in order to provide behavioral anxiety indicators during speech. These include scoring and analyzing behaviors such as speech dysfluency, positioning and movements for hands/arms, and position of the body and body language.

Experimental Design and Procedure

Upon recruitment, participants will be asked if they agree with participating in the experiment through electronic informed consent and an online questionnaire that includes questions regarding demographic information, the PRCS-12, and the LSAS. Those who score above the cut-off points will qualify for the experiment. Those qualified will then be invited to register for a pre-arranged timeslot for session 1 of the experiment. Upon arrival, participants will be required to complete the IAT and the SPIC. After completion, participants will register for a timeslot 1 week later for VR exposure. The first VR session will follow the same procedure as Experiment 1, except for the introduction of splitting variations and SUDS questions. Participants will be randomly assigned to one of the four experimental conditions or the control condition. The experimental conditions will be based on the proximity and transparency of the therapist splitting. This creates a between-groups design with 2 factors, "Transparency" with levels Opaque and Transparent, and "Proximity" with levels Near and Far. Therefore, the groups are designated as so: Near+Opaque, Near+Transparent, Far+Opaque, and Far+Transparent (Figure 4). Additionally, a within-group factor of Session (initial, VR, and follow-up) will also be incorporated.

Once assigned, participants will be immersed in the IVRE through the HMD. In both groups, participants will be embodied in a same-gender virtual body that will move synchronously with the participant's real movements. Once in the environment, the participants will find themselves talking to Sigmund Freud, the same virtual therapist from Experiment 1. The virtual therapist will be programmed to talk with the participant about their fear of public speaking, and it will be identical to Experiment 1 for both groups. Participants will also be asked to participate in some public speaking exercises, including describing a hobby or the synopsis of a book, movie, or television show. During the exercises, particularly when the participants are engaging for an extended period of time, the therapist will begin to divide and copy himself, repeatedly, until eventually a group of copies has appeared in the environment. The splitting variations occur here, in which the therapist will appear to split in an opaque or transparent manner. After completion of the split, the copy will either proceed to move directly to the side of the room (Near) or turn and walk away from the splitting location, ultimately residing outside of the personal space (Far). Those copies of the therapist will also morph and change into different, random human characters after a period of time, whom eventually take seats in the chairs behind the therapist. Immediately following the session, participants will provide responses for the IAT, SPIC, and VRQ.

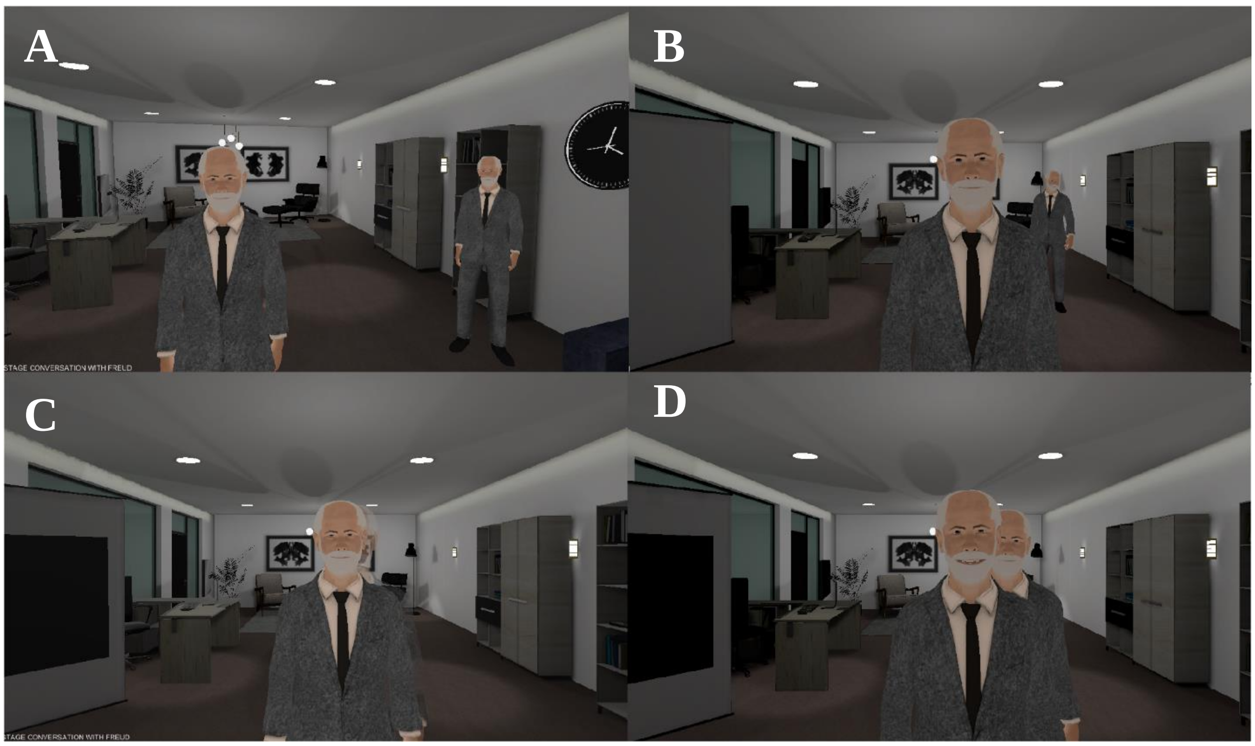

Figure 4. Experimental conditions: The splitting methods are provided, with the top row demonstrating Proximity (**A**: Near/**B**: Far) and the bottom row demonstrating Transparency (**C**: Transparent/**D**: Opaque).

In addition, a control group will also be employed, in which individuals will experience 6 sequential sessions of exposure therapy to simulate in vivo methods. The first session will consist of only interaction and dialogue with the virtual therapist to establish a therapeutic relationship and to provide psychoeducation to the participant. At the end of the first session (and each subsequent session), the therapist will inform the participant of the number of avatars that will be present for the next session. The following session will occur 2 to 3 days later, in which participants will engage in similar public speaking exercises as the experimental conditions. Each sequential session following the initial will also consist of public speaking exercises in front of an increasing number of virtual audience avatars, beginning with 3 and progressing to 15 by groups of 3. After the end of the last session, participants will provide their responses for the IAT, SPIC, and VRQ and sign-up for the final follow-up session.

Approximately one week later, all participants will return to give a speech to a virtual audience. Prior to immersion, participants will be allotted 10 minutes to prepare their speech, which will consist of introducing a fabricated band and providing recognition for their various achievements. Experimental data will also be collected prior to and following this session, in order to assess participants' anxiety (IAT) and perceived competence (SPIC) levels before and after the virtual experience. In addition, once participants are finished, they will exit the IVRE and record their responses for the VRQ. The participants will then be debriefed on the study and asked to experience the other experimental conditions, up until after the point of splitting. The control condition will also experience all splitting methods. Experience of the other conditions will be recorded in the post-VR interview. After the interview, participants will be compensated for their complete participation.

Proposed Data Analysis

Descriptive statistics and preliminary analyses will be conducted for all response variables, particularly to identify if violations are present for the assumptions of normality, homoscedasticity, and outliers.

Given the provided hypotheses for this experiment, mixed-factor ANOVAs will be conducted on the response variables of the IAT and SPIC to determine if differences are present between the four experimental conditions and the control condition (between) and also between the sessions (within). In the case that multiple comparisons are needed, independent samples and repeated-measures T-tests will be conducted in a pairwise manner for the IAT and SPIC scores. Provided that the VRQ is scored on an ordinal scale and tends towards non-normal distributions, non-parametric Kruskal-Wallis H-tests will be conducted on the VRQ to determine differences between the experimental conditions and the control. Further pairwise Mann-Whitney U-tests will be utilized for multiple comparisons of the groups. In addition, a factor analysis of the VRQ will be executed on the responses in an attempt to validate the results from Experiment 1.

Due to the PRCS-12 and the LSAS being utilized as screening measures, linear regressions will be employed to determine if either acts as a predictive measure for the change in scores for the SPIC and IAT, respectively. SUDS scores will be also be analyzed with a mixed-factor ANOVA, investigating the change in score over sessions (within) and the change in score between conditions (between). Analysis of the post-VR interview will be conducted to determine the presence of a splitting preference. This will be done through comparison of coded and grouped responses converted from the interview. No hypotheses have been provided for these measures. Finally, the behavioral measures will be included in the analysis as a means to explore anxious behavior during speech and the relation between behavior and subjective measures.

General Discussion

Experiment 1 is the first study to implement a splitting method in an IVRE that would simulate gradual exposure to an audience in a single session. While research has shown the use of multi-session exposure therapy in VR as a treatment method for PSA, focusing on therapist-led CBT interventions[4,45] and self-led multi-session exposure therapy[1,9,47], none have attempted self-led single-session rapid exposure therapy. Interestingly, the therapist-led and self-led therapeutic approaches have both demonstrated positive changes for PSA, measured through marked decreases in SUDS, PRCS, LSAS, and BFNE, to only name a few. The results of Experiment 1 are however rather inconclusive, as it was observed that the control and experimental conditions did not display many significant differences. Trends in the analyses demonstrated negative changes for the BFNE and the IAT across sessions for the control condition, while the experimental condition tended to show negative changes for the STAI and SPIC. Those in the control condition experienced greater negative evaluation after giving a speech to an audience after only speaking with a therapist, with a slightly medium effect size for the BFNE (Cohen's $d = 0.473$) and a large effect size for IAT (Cohen's $d = 0.813$) were produced. The IAT has been proven to be an effective predictor for behavioral anxiety indicators, demonstrating lower scores on the IAT were significantly associated with more nervous movements of the mouth and hands and speech dysfluency[48]. Therefore, use of the IAT in Experiment 2, in conjunction with the behavioral analyses of the recorded speeches, may provide supporting evidence for the aforementioned results.

On the other hand, the experimental condition displayed a greater, negative change in SPIC scores compared to the control condition (Cohen's d = 0.872). Interestingly, the SPIC score was based on the time between sessions, which would therefore indicate those in the experimental condition perceived themselves as less competent to give the speech in the follow-up. If this is true, it could be possible that participants overestimated their ability to give a speech, prior to the VR session. Participation in the experimental condition, in which the crowd appears, could give rise to individuals updating their prior beliefs and subsequently feeling less competent towards giving the next speech in the second session. An attempt to test for this is to be implemented in Experiment 2 by measuring the SPIC at four timepoints, rather than only two, pre- and post-sessions 1 and 2. This could provide a better indication as to why SPIC scores would change over time. Additionally, the experimental condition did display similar results on the BFNE compared to a study that examined single-session self-led VRET compared to a wait-list condition. In the study, no significant changes in BFNE scores were observed between the two conditions on six- and twelve-month follow-ups[4]. This could be directly attributed to the fact the audience is not real and therefore participants do not feel as if they will be negatively evaluated by the virtual avatars. However, other results have shown that CBT paired with VR exposure elicits an equal decrease in BFNE scores as CBT paired with in vivo exposure for social anxiety[49]. Therefore, it may be that the BFNE does not target fear of public speaking and PSA as it does for social anxiety disorder.

The factor analysis from Experiment 1 discovered the factor consisting of *VRBody* and *Agency* demonstrated a lack of differences between conditions and sessions, helping to ensure a consistent degree of immersion in the IVRE for all participants, even as their virtual bodies did not directly resemble their own. This is in line with other studies that show it is possible to report a high degree of body ownership and agency when inhabiting a different body[9,17,34]. As previously stated, it is important to demonstrate that the sample as a whole possess high illusions of body ownership and agency because the experiment is reliant upon the fact that participants will behave and respond in a realistic manner. Additionally, the first factor, Y_{aud}, provided strong evidence that both conditions responded as if they were speaking to a real audience in the follow-up in relation to the first session ($\eta^2 = 0.586$). However, with respect to the experimental condition that was rapidly exposed to an audience in the first session, the factor scores did not reflect that. In Experiment 2, this will also specifically be addressed through the inclusion of the post-VR interview. The rationale for including the post-VR interview is to not only extract qualitative responses regarding the experiment, but also quantitative ones regarding the audience that can directly attribute to strengthening the validity of the IVRE and experimental paradigm.

The study does possess moderate limitations, some of which have already been addressed. In addition to those, the size of the sample leads to doubts regarding the generalizability of the results to the greater population. Results from the statistical analyses indicated low levels of power, which can be directly attributed to the small sample. Moreover, the overall results from the sample comprised mostly of low scores, relative to each measure. This would indicate that the sample was not entirely representative of the target population, who possess high degrees of PSA and therefore could potentially be more responsive to the paradigm and dependent measures. Experiment 2 will ideally address the limitations that have been provided.

Future studies should aim to explore any method that can simulate the progression of an audience. What may be of interest for the future is to examine participants' eye gaze during the splitting and morphing processes. Since a significant portion of the aim of both experiments is to determine how participants react to these processes, it may be beneficial to determine where participants are directing their gaze and attention during the experiment, and especially while these processes are occurring. Limited research is available regarding eye tracking in VR under the intention of studying the perceptibility of scene changes, and nearly none regarding scene changes that involve a change in the number of avatars. Some recent and related work examined saccadic movements during scene transformations and determined that slight changes to the point-of-view of the participant during a vertical or horizontal saccade, caused by a visual stimulus, were highly imperceptible. When given a forced-choice task to tell which direction the scene shifted, participants responded with higher accuracy to rotational changes versus translational changes.[50] This would fall in line with concepts in perception and attention in which this experiment incorporated attentional capture and change blindness to alter the scene imperceptibly to the participant. Part of the rationale for having the therapist begin the splitting process while a participant is engaged in a public speaking exercise is based on attentional capture, so that the participant's attention is on the act of speaking and completing the exercise. Eye tracking data could be analyzed alongside the SUDS scores to determine if attention on the emerging crowd is associated with changes in self-reported distress. It is also the goal of future studies to continue to gather convincing evidence that IVREs can elicit natural, realistic responses from users, even when the environment is manipulated in a manner that violates reality. Taken all together, such evidence helps to push forward the possibility that rapid exposure for VRET is effective and can be utilized as a means to prepare individuals for public speaking. However, much more work on this will need to be completed before firm conclusions can be made.

References

1. Takac, M. *et al.* Public speaking anxiety decreases within repeated virtual reality training sessions. *PLoS ONE* **14**, e0216288 (2019).
2. Pertaub, D.-P., Slater, M. & Barker, C. An experiment on fear of public speaking in virtual reality: (705412011-025). (2001) doi:10.1037/e705412011-025.
3. Zacarin, M. R. J. *et al.* Behavioral Therapy and Virtual Reality Exposure for Public Speaking Anxiety. *Trends in Psychology* **27**, 491–507 (2019).
4. Lindner, P. *et al.* Therapist-led and self-led one-session virtual reality exposure therapy for public speaking anxiety with consumer hardware and software: A randomized controlled trial. *Journal of Anxiety Disorders* **61**, 45–54 (2019).
5. Ruscio, A. M., Brown, T. A., Chiu, W. T., Sareen, J. & Kessler, R. C. Social Fears and Social Phobia in the United States: Results from the National Comorbidity Survey Replication. 20 (2008).
6. Mesri, B. *et al.* Public speaking avoidance as a treatment moderator for social anxiety disorder. *Journal of Behavior Therapy and Experimental Psychiatry* **55**, 66–72 (2017).
7. Levin, M. E., Haeger, J. & Smith, G. S. Examining the Role of Implicit Emotional Judgments in Social Anxiety and Experiential Avoidance. *J Psychopathol Behav Assess* **39**, 264–278 (2017).
8. Feske, U. & Chambless, D. L. Cognitive behavioral versus exposure only treatment for social phobia: A meta-analysis. *Behavior Therapy* **26**, 695–720 (1995).
9. Slater, M. *et al.* An experimental study of a virtual reality counselling paradigm using embodied self-dialogue. *Sci Rep* **9**, 10903 (2019).
10. Landowska, A., Roberts, D., Eachus, P. & Barrett, A. Within- and Between-Session Prefrontal Cortex Response to Virtual Reality Exposure Therapy for Acrophobia. *Front. Hum. Neurosci.* **12**, (2018).
11. Slater, M. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Phil. Trans. R. Soc. B* **364**, 3549–3557 (2009).
12. Skarbez, R., Neyret, S., Brooks, F. P., Slater, M. & Whitton, M. C. A Psychophysical Experiment Regarding Components of the Plausibility Illusion. *IEEE Trans. Visual. Comput. Graphics* **23**, 1369–1378 (2017).

13. Botvinick, M. & Cohen, J. Rubber hands ‘feel’ touch that eyes see. *Nature* **391**, 756–756 (1998).
14. Ehrsson, H. H., Wiech, K., Weiskopf, N., Dolan, R. J. & Passingham, R. E. Threatening a rubber hand that you feel is yours elicits a cortical anxiety response. *Proceedings of the National Academy of Sciences* **104**, 9828–9833 (2007).
15. Sanchez-Vives, M. V., Spanlang, B., Frisoli, A., Bergamasco, M. & Slater, M. Virtual Hand Illusion Induced by Visuomotor Correlations. *PLOS ONE* **5**, e10381 (2010).
16. Slater, M., Pérez Marcos, D., Ehrsson, H. & Sanchez-Vives, M. V. Towards a digital body: the virtual arm illusion. *Front. Hum. Neurosci.* **2**, (2008).
17. Osimo, S. A., Pizarro, R., Spanlang, B. & Slater, M. Conversations between self and self as Sigmund Freud—A virtual body ownership paradigm for self counselling. *Sci Rep* **5**, 13899 (2015).
18. Hall, E. T. *The Hidden Dimension*. (Doubleday, 1966).
19. Wilcox, L. M., Allison, R. S., Elfassy, S. & Grelik, C. Personal space in virtual reality. *ACM Transactions on Applied Perception* **3**, 17 (2006).
20. Bailenson, J. N., Blascovich, J., Beall, A. C. & Loomis, J. M. Interpersonal Distance in Immersive Virtual Environments. *Pers Soc Psychol Bull* **29**, 819–833 (2003).
21. Llobera, J., Spanlang, B., Ruffini, G. & Slater, M. Proxemics with multiple dynamic characters in an immersive virtual environment. (2010).
22. Mostajeran, F. & Hamburg, U. The Effects of Virtual Audience Size on Social Anxiety During Public Speaking. 10 (2020).
23. Comer, J. S. & Kendall, P. C. *The Oxford Handbook of Research Strategies for Clinical Psychology*. (OUP USA, 2013).
24. Paul, G. L. *Insight Vs. Desensitization in Psychotherapy: An Experiment in Anxiety Reduction*. (Stanford University Press, 1966).
25. Hook, J. N., Smith, C. A. & Valentiner, D. P. A short-form of the Personal Report of Confidence as a Speaker. *Personality and Individual Differences* **44**, 1306–1313 (2008).
26. Collins, K. A., Westra, H. A., Dozois, D. J. A. & Stewart, S. H. The validity of the brief version of the Fear of Negative Evaluation Scale. *Journal of Anxiety Disorders* **19**, 345–359 (2005).

27. Greenwald, A. G., McGhee, D. E. & Schwartz, J. L. K. Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology* **74**, 1464–1480 (1998).
28. Greenwald, A. G. & Farnham, S. D. Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology* **79**, 1022–1038 (2000).
29. JULIAN, L. J. Measures of Anxiety. *Arthritis Care Res (Hoboken)* **63**, (2011).
30. Tluczek, A., Henriques, J. B. & Brown, R. L. Support for the reliability and validity of a six-item state anxiety scale derived from the State-Trait Anxiety Inventory. *J Nurs Meas* **17**, 19–28 (2009).
31. MacIntyre, P. D. & Renée MacDonald, J. Public speaking anxiety: Perceived competence and audience congeniality. *Communication Education* **47**, 359–365 (1998).
32. Aymerich-Franch, L. & Baileson, J. The use of doppelgangers in virtual reality to treat public speaking anxiety: a gender comparison. 15 (2014).
33. Banakou, D. & Slater, M. Body ownership causes illusory self-attribution of speaking and influences subsequent real speaking. *Proc Natl Acad Sci USA* **111**, 17678–17683 (2014).
34. Banakou, D., Kishore, S. & Slater, M. Virtually Being Einstein Results in an Improvement in Cognitive Task Performance and a Decrease in Age Bias. *Front. Psychol.* **9**, 917 (2018).
35. McKinney, W. Data Structures for Statistical Computing in Python. in 56–61 (2010). doi:10.25080/Majora-92bf1922-00a.
36. Virtanen, P. *et al.* SciPy 1.0--Fundamental Algorithms for Scientific Computing in Python. *Nat Methods* **17**, 261–272 (2020).
37. Vallat, R. Pingouin: statistics in Python. *Journal of Open Source Software* **3**, 1026 (2018).
38. Seabold, S. & Perktold, J. Statsmodels: Econometric and Statistical Modeling with Python. in 92–96 (2010). doi:10.25080/Majora-92bf1922-011.
39. Michael Waskom *et al. mwaskom/seaborn: v0.10.1 (April 2020)*. (Zenodo, 2020). doi:10.5281/zenodo.3767070.
40. Heimberg, R. G. *et al.* Psychometric properties of the Liebowitz Social Anxiety Scale. *Psychological Medicine* **29**, 199–212 (1999).
41. Rytwinski, N. K. *et al.* Screening for social anxiety disorder with the self-report version of the Liebowitz Social Anxiety Scale. *Depression and Anxiety* **26**, 34–38 (2009).

42. VR device shipments by vendor worldwide 2017-2019. *Statista* https://www.statista.com/statistics/671403/global-virtual-reality-device-shipments-by-vendor/.
43. Wolpe, J. & Lazarus, A. A. *Behavior therapy techniques: A guide to the treatment of neuroses*. ix, 198 (Pergamon Press, 1966).
44. Benjamin, C. L. *et al.* Patterns and Predictors of Subjective Units of Distress in Anxious Youth. *Behav Cogn Psychother* **38**, 497–504 (2010).
45. Harris, S. R., Kemmerling, R. L. & North, M. M. Brief Virtual Reality Therapy for Public Speaking Anxiety. *CyberPsychology & Behavior* **5**, 543–550 (2002).
46. Rauch, S. A. M. *et al.* In session extinction and outcome in Virtual Reality Exposure Therapy for PTSD. *Behaviour Research and Therapy* **109**, 1–9 (2018).
47. Kählke, F. *et al.* Efficacy of an unguided internet-based self-help intervention for social anxiety disorder in university students: A randomized controlled trial. *International Journal of Methods in Psychiatric Research* **28**, e1766 (2019).
48. Egloff, B. & Schmukle, S. C. Predictive validity of an implicit association test for assessing anxiety. *Journal of Personality and Social Psychology* **83**, 1441–1455 (2002).
49. Bouchard, S. *et al.* Virtual reality compared with in vivo exposure in the treatment of social anxiety disorder: A three-arm randomised controlled trial. *The British Journal of Psychiatry* **210**, 276–283 (2017).
50. Keyvanara, M. & Allison, R. Transsaccadic Awareness of Scene Transformations in a 3D Virtual Environment. in *ACM Symposium on Applied Perception 2019* 1–9 (Association for Computing Machinery, 2019). doi:10.1145/3343036.3343121.

Appendix

	Condition	
	Freud Crowd	**Freud Only**
Female (Male)	n = 3 (1)	n = 4 (1)
Age Mean ± SD	22.0 (3.08)	21.5 (2.65)
Literacy Median (IQR)	5 (0)	4.5 (1)
Programming Median (IQR)	2 (1)	2 (1.25)
VR Experience Median (IQR)	3 (4)	3 (2.25)
Games Median (IQR)	3 (4)	3 (3)
Gametime Median (IQR)	1 (1)	1 (0.5)

Table A. Descriptive statistics for the demographic questionnaire.

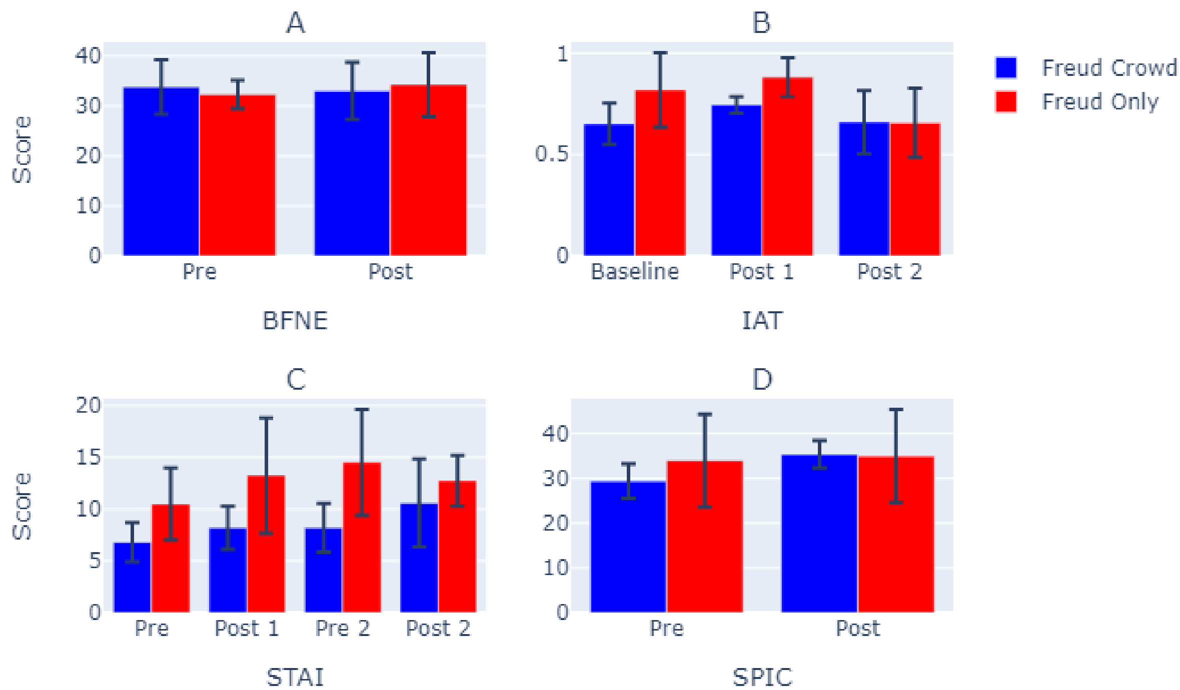

Figure A. Bar charts displaying the raw mean scores and standard errors for each measurement: (**A**) BFNE, (**B**) IAT, (**C**) STAI, (**D**) SPIC.

Negative Evaluation

Brief Fear of Negative Evaluation

	T	tail	p	CI95%	cohen-d	BF10	power
T-test	-0.669	one-sided	0.267	[-inf, 5.67]	0.473	1.164	0.157

Table B. T-test for *dBFNE*.

Implicit Association Test

Source	SS	DF1	DF2	MS	F	p	n2	eps
condition	0.045	1	7	0.045	0.478	0.512	0.034	NaN
Scale	0.097	1	7	0.097	1.294	0.293	0.072	1.0
Interaction	0.022	1	7	0.022	0.291	0.607	0.016	NaN

Table C. Mixed-factorial ANOVA for *dIAT*. Overall model for *dIAT*: $F(3,14) = 0.644$, p = 0.599, $r^2 = 0.121$.

Anxiety

State Trait Anxiety Inventory

Source	SS	DF1	DF2	MS	F	p	n2	eps
condition	0.044	1	7	0.044	0.002	0.968	0.000	NaN
Scale	5.556	1	7	5.556	0.185	0.680	0.014	1.0
Interaction	9.344	1	7	9.344	0.311	0.594	0.023	NaN

Table D. Mixed-factorial ANOVA for *dSTAI*. Overall model for *dSTAI*: $F(3,14) = 0.179$, p = 0.909, $r^2 = 0.037$.

Speaker Competence

Personal Report of Confidence as Speaker

	T	tail	p	CI95%	cohen-d	BF10	power
T-test	0.406	two-sided	0.697	[-15.28, 21.58]	0.268	0.535	0.064

Table E. T-test for the PRCS.

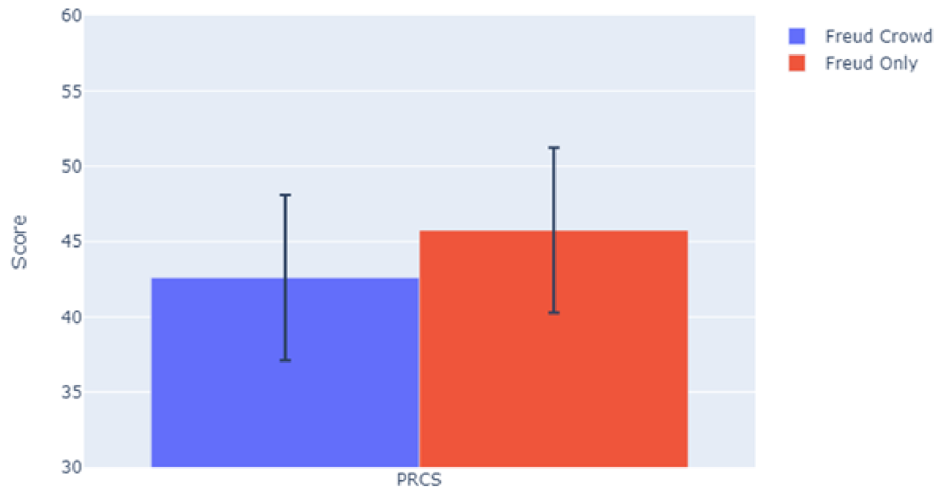

Figure B. Bar chart of means and standard errors for PRCS, showing the scores taken before the VR experience.

State Perceived Index of Competence

	T	tail	p	CI95%	cohen-d	BF10	power
T-test	1.403	two-sided	0.209	[-3.69, 13.69]	0.872	0.875	0.204

Table F. T-test for *dSPIC*.

Virtual Reality

Virtual Reality Questionnaire

Variable	Question	Variable (n)
VRBody	I felt that the virtual body I saw when looking down at myself was my own body.	vrbody (n=9) vrbody2 (n=9)
Mirror	I felt that the virtual body I saw when looking at myself in the mirror was my own body	mirror (n=9) mirror2 (n=9)
Agency	During the experience there were moments that I felt that if I moved my real body (hands, arms, legs), my virtual body would move equivalently.	agency (n=9) agency2 (n=9)
Relaxed	How comfortable and relaxed did you feel during your virtual experience?	relaxed (n=9) relaxed2 (n=9)
Presence	How much did you have the feeling that there was an audience in front of you?	presence (n=9) presence2 (n=9)
Audience	How conscious were you of having an audience in front of you?	audience (n=9) audience2 (n=9)
Impression	What impression do you believe you gave the audience?	impression (n=9) impression2 (n=9)
Pleasant	How pleasant was the audience?	pleasant (n=9) pleasant2 (n=9)

Table G. VRQ variables and descriptions. Variables marked with a 2 were measured following session 2. The preceding variables were measured following session 1.

Contrast	Question	A	B	U-val	Tail	p	Cohen's d
Question * condition	vrbody	Freud Crowd	Freud Only	5.0	two-sided	0.218	0.894
Question * condition	mirror	Freud Crowd	Freud Only	9.5	two-sided	1.000	0.048
Question * condition	agency	Freud Crowd	Freud Only	5.0	two-sided	0.241	-0.433
Question * condition	relaxed	Freud Crowd	Freud Only	6.0	two-sided	0.379	-0.661
Question * condition	presence	Freud Crowd	Freud Only	10.5	two-sided	1.000	0.411
Question * condition	audience	Freud Crowd	Freud Only	9.0	two-sided	0.901	-0.023
Question * condition	impression	Freud Crowd	Freud Only	6.0	two-sided	0.308	-1.011
Question * condition	pleasant	Freud Crowd	Freud Only	5.5	two-sided	0.297	-0.887
Question * condition	vrbody2	Freud Crowd	Freud Only	6.0	two-sided	0.338	-0.764
Question * condition	mirror2	Freud Crowd	Freud Only	10.0	two-sided	0.867	0.199
Question * condition	agency2	Freud Crowd	Freud Only	4.0	two-sided	0.152	-0.956
Question * condition	relaxed2	Freud Crowd	Freud Only	8.5	two-sided	0.802	-0.148
Question * condition	presence2	Freud Crowd	Freud Only	9.5	two-sided	1.000	0.048
Question * condition	audience2	Freud Crowd	Freud Only	9.5	two-sided	1.000	-0.203
Question * condition	impression2	Freud Crowd	Freud Only	5.0	two-sided	0.247	-1.016
Question * condition	pleasant2	Freud Crowd	Freud Only	11.5	two-sided	0.793	0.202

Table H. Table displaying the pairwise Mann-Whitney U tests between the two experimental conditions for sessions 1 and 2 for the individual variables for the VRQ.

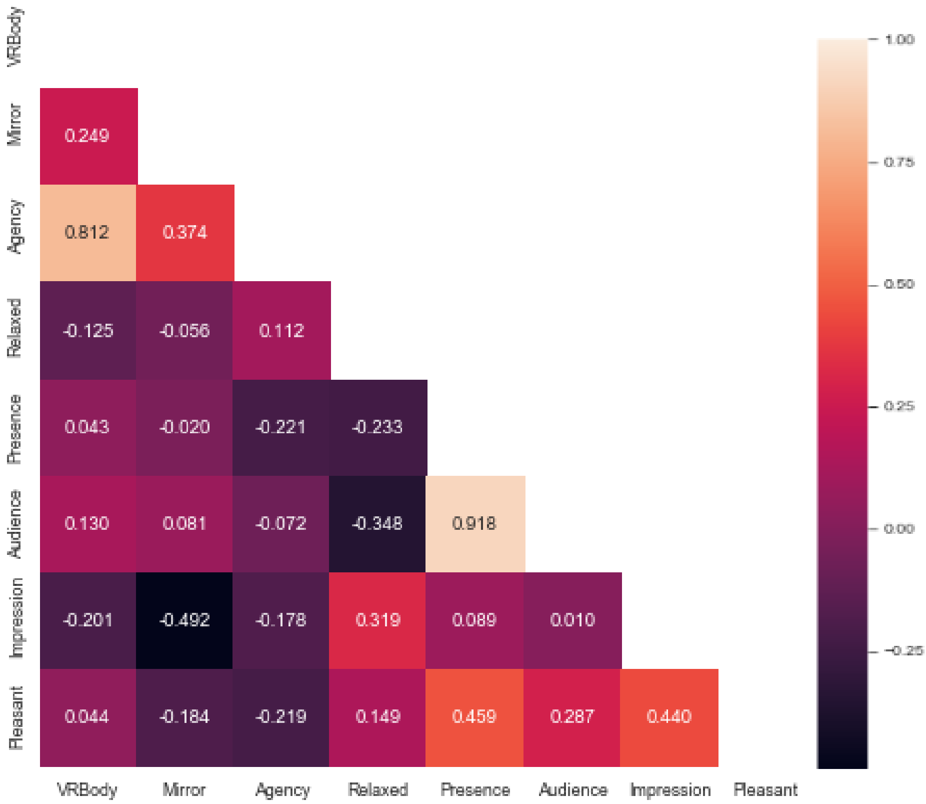

Figure C. Spearman correlation matrix between the raw variable scores of the VRQ. The color bar indicates that stronger correlations are depicted with lighter or darker colors at the extremes.

Variable	***F1***	***F2***	**uniqueness**
VRBody	-0.0121	0.9875	0.0061
Mirror	0.0440	04951	0.7409
Agency	-0.1824	0.8162	0.3283
Relaxed	-0.3941	-0.2497	0.7467
Presence	0.9842	-0.1244	0.0200
Audience	0.9135	-0.0542	0.1499
Impression	0.0216	-0.4241	0.8183
Pleasant	0.4103	-0.0973	0.8283

Table I. Factor loadings and uniqueness from the factor analysis for the VRQ.

Variable	Y_{aud} (F1)	Y_{body} (F2)
VRBody	0.2143	0.8908***
Mirror	0.0789	0.3463
Agency	-0.0794	0.9200***
Relaxed	-0.3202	0.0042
Presence	0.9589***	-0.2325
Audience	0.9541***	-0.0673
Impression	0.0328	-0.2646
Pleasant	0.3912	-0.1463

Table J. Spearman correlations between the factors and original variables for the VRQ. *** $p < 0.001$.

Source	SS	DF	MS	F	p-unc	n2	eps
Session	4.2982	1	4.2982	11.3383	0.0098	0.5863	1.0
Error	3.0327	8	0.3791	NaN	NaN	NaN	NaN

Table K. Repeated-measures one-way ANOVA for *F1* of the factor analysis.

Source	SS	DF	MS	F	p-unc	n2	eps
Session	0.271	1	0.271	0.359	0.566	0.043	1.0
Error	6.044	8	0.756	NaN	NaN	NaN	NaN

Table L. Repeated-measures one-way ANOVA for *F2* of the factor analysis.

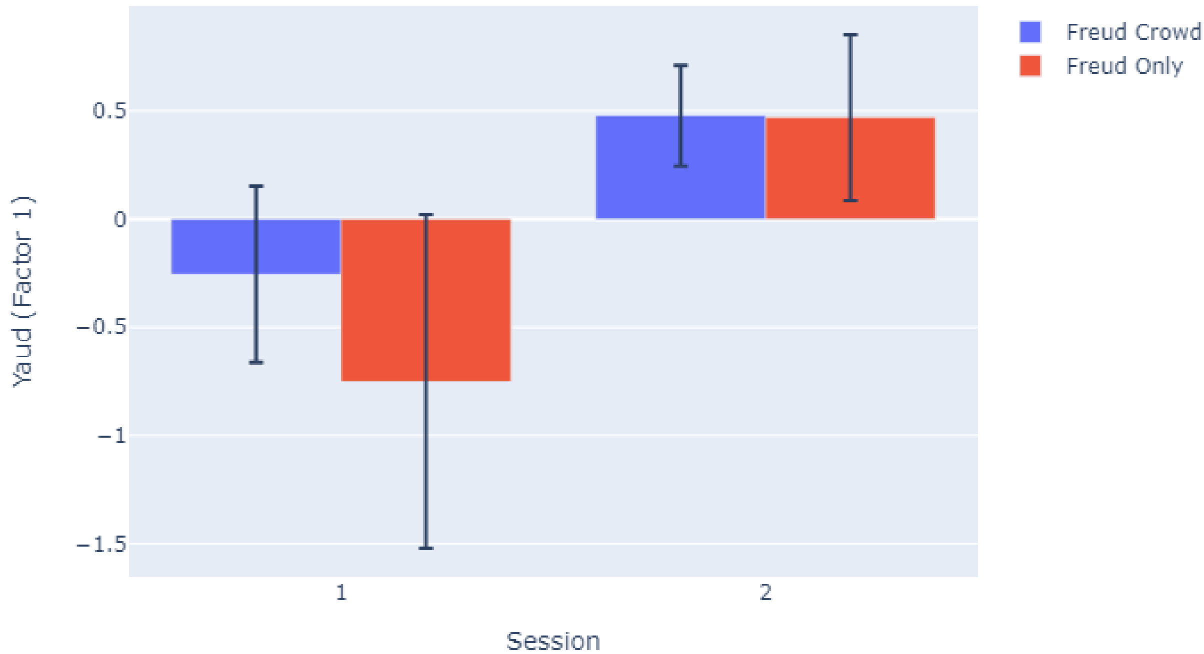

Figure D. Bar chart with means and standard errors for Factor 1 from the factor analysis. Factor 1 consists of *Presence* and *Audience*.

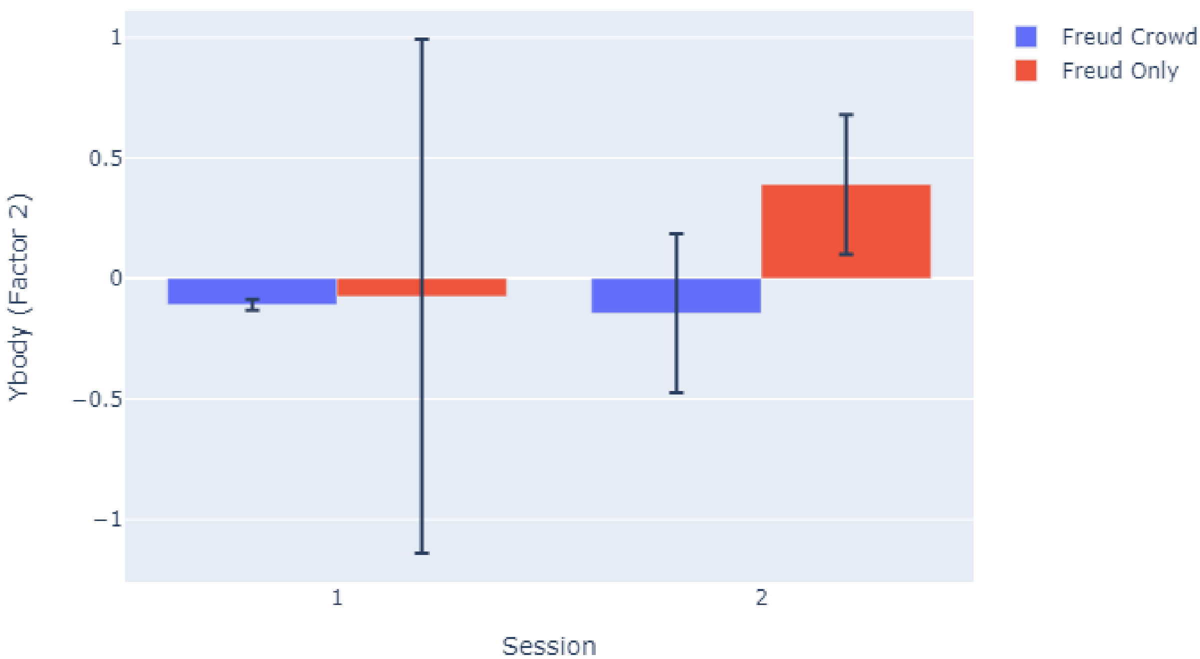

Figure E. Bar chart with means and standard errors for Factor 2 from the factor analysis. Factor 2 consists of *VRBody* and *Agency*.

Post-VR Interview Questions

1. At what point during the process was there a presence of a crowd?
2. How aware were you (the participant) that a crowd was beginning to fill the environment?
3. Did the process of splitting and morphing lead to the feeling of being in the presence of an audience/crowd?
4. When is the fear/anxiety of public speaking present?
5. Which method appeared to cause the least amount of anxiety? The most?
6. Which method was the least/most distracting?
7. How distracting was the splitting process?
8. When the splitting process began, was focus lost from the task at hand?
9. Did the splitting/morphing process feel natural in the environment? Or artificial?
10. Which splitting method was most preferred?