Pilot randomized trial of self-guided virtual reality exposure therapy for social anxiety disorder

Nur Hani Zainal[a,*], **William W. Chan**[b], **Alisha P. Saxena**[b], **Craig Barr Taylor**[b,c], **Michelle G. Newman**[a]

[a]The Pennsylvania State University, USA

[b]Palo Alto University, USA

[c]Stanford University, USA

Abstract

Background: Virtual reality exposure therapy (VRE) has shown promising efficacy for the treatment of social anxiety disorder (SAD) and related comorbidities. However, most trials conducted to date were therapist-led, and little is known about the efficacy of *self-guided* VRE. Therefore, this randomized controlled trial (RCT) aimed to determine the efficacy of a self-directed VRE for SAD.

Method: Forty-four community-dwelling or undergraduate adults diagnosed with SAD based on the Mini International Neuropsychiatric Interview were randomly assigned to VRE designed to last four sessions or more ($n = 26$) or waitlist (WL; $n = 18$). Self-reported SAD severity (Social Phobia Diagnostic Questionnaire and Social Interaction Anxiety Scale), job interview anxiety (Measure of Anxiety in Selection Interviews), trait worry (Penn State Worry Questionnaire), and depression symptoms (Patient Health Questionnaire-9) were administered at baseline, post-

[*]Corresponding author. Department of Psychology, 378 Moore Building, University Park, PA, 16802, USA, nvz5057@psu.edu (N.H. Zainal).

treatment, 3-month-follow-up (3MFU), and 6-month-follow-up (6MFU). Piecewise multilevel modeling analyses were conducted to manage clustering in the data.

Results: VRE vs. WL resulted in greater reductions in SAD symptom severity, job interview fear, and trait worry, with moderate-to-large effect sizes (Hedge's $g = -0.54$ to -1.11) from pre-to-post treatment. Although significant between-group differences did not emerge for change in depression, VRE led to change in depression, whereas waitlist did not. These gains were also maintained at 3MFU and 6MFU. Further, facets of presence increased during the course of VRE ($g = 0.36–0.45$), whereas cybersickness decreased ($g = -0.43$).

Discussion: Brief, self-guided VRE might ameliorate SAD and comorbid worry, for young-to-middle-aged adults with SAD. Other theoretical and practical implications were also discussed.

Keywords

Social anxiety disorder (SAD) is a widespread mental health problem (Kessler et al., 2012), marked by persistent and intense fear and avoidance of interpersonal situations that may be evaluative (American Psychiatric Association, 2013). SAD is acknowledged to be a complex condition that can be chronic or unremitting (Vriends, Bolt, & Kunz, 2014). Relatedly, persons with SAD tend to display compromised work or school functioning, relationship satisfaction, and quality of life (Teo, Lerrigo, & Rogers, 2013), as well as increased comorbid physical and mental health problems (Jacobson & Newman, 2017; Schmidt, Keough, Hunter, & Funk, 2008). Therefore, designing and delivering efficacious treatments for SAD is important.

Novel state-of-the-art psychotherapies for SAD have been developed. These include cognitive bias modification approaches (e.g., video feedback, imagery rescripting), cognitive behavioral therapies (CBT; e. g., social skills training, behavioral experiments), internet CBT, and neuroenhancers (e.g., D-cycloserine, yohimbine) (Stangier, 2016). Exposure therapy is a well-established, gold-standard efficacious specific CBT component for SAD (Otte, 2011). Such exposure therapy is grounded in *emotional processing theory* (Foa & Kozak, 1986), which proposes that fear memories can be interpreted as structures that store data concerning events, responses, and meaning propositions. By activating and altering fear structures, it is thought that exposure therapy can ease emotional processing through gradual and recurrent experiences of distress-evoking yet potentially rewarding and physically safe social interactions (Zalta & Foa, 2012, pp. 75–104). Specifically, exposure therapy is supposed to activate the fear structure by involving the person with SAD in various anxiety-provoking social situations to adjust the fear structure via habituation (reduced fear after repeated encounters of phobic situations) and extinction (unlearning of acquired fear connections; Furlong, Richardson, & McNally, 2016). On the other hand, it is plausible that even if within- or between-session habituation did not occur, extinction learning might be achieved via practicing greater fear tolerance in diverse anxiety-evoking social situations (cf. *inhibitory learning theory*; Craske et al., 2008). Relatedly, the *cognitive model* of SAD (Leigh & Clark, 2018) posits that exposure therapy works by improving psychological

flexibility and other (vs. self-) focused attention as well as eliminating safety behaviors and threat hypervigilance, during anxiety-inducing social events.

Exposure therapy can be carried out *in vivo* (directly engaging with fear-evoking situations) or *in vitro* or imaginarily (continually visualizing exposure scenarios). However, *in vivo* exposure therapy can be time-consuming, costly, and how events unfold during exposure can neither be predicted nor controlled. Also, *in vitro* exposure therapy can be difficult and ineffective for highly-avoidant persons who lack the ability to envision specific scenes in clear detail and to stay in the imagined anxiety-inducing scenarios long enough for the imaginal exposure therapy to be effective (Hembree & Cahill, 2007; Vrielynck & Philippot, 2009). Thus, virtual reality exposure therapy (VRE) has been designed as a means for persons with SAD to privately immerse in myriad computer-generated fear-provoking social situations that mimic the natural environment (Anderson, Rothbaum, & Hodges, 2003). VRE offers latitude for the psychologist to construct exposure situations, receives more acceptance from patients, provides greater privacy, and has lower dropout rates than *in vivo* exposure therapy (Garcia-Palacios, Botella, Hoffman, & Fabregat, 2007; Krijn, Emmelkamp, Olafsson, & Biemond, 2004). Further, most clinicians have viewed VRE favorably (Lindner, Miloff, Zetterlund, et al., 2019).

To date, seven randomized controlled trials (RCTs) have examined the impact of therapist-directed VRE compared to waiting list (WL) for SAD. An early study showed that four therapist-facilitated 12-to-15-min sessions of VRE (vs. WL) efficaciously reduced public speaking phobia for highly anxious college students from baseline to post-treatment (Harris, Kemmerling, & North, 2002); despite this reduction, VRE was not superior to WL in lowering SAD symptoms. Another four-session VRE (vs. WL) that instructed participants to repeatedly read a children's book to a virtual audience observed reductions in public speaking fear across time (Lister, Piercey, & Joordens, 2010); however, there were no changes on trait anxiety or social phobia symptoms. Relatedly, compared to WL, a therapist-guided, 12-session, VRE with various fear-inducing interpersonal situations led to decrements in SAD fear and avoidance severity, as well as public speaking phobia, among young adult females with public speaking anxiety (Wallach, Safir, & Bar-Zvi, 2011). Further, some sustained treatment gains at 1-year follow-up were observed (Safir, Wallach, & Bar-Zvi, 2012); however, high-anxious participants were not assessed for diagnoses of SAD in that study. In a similar nine-session, therapist-directed VRE (vs. WL) RCT for public speaking anxiety that additionally recruited patients with SAD and assigned between-session homework, large effect sizes for treatment efficacy were observed from baseline to 1-year follow-up (Anderson et al., 2013); nonetheless, findings may not extend beyond persons who had expressed public speaking fear as their main problem. Relatedly, VRE (vs. WL) was more effective in reducing avoidance and unhelpful beliefs as well as increasing speech duration in persons with SAD whose interpersonal fears extended to numerous contexts (Kampmann et al., 2016). Additionally, cognitive behavioral therapy (CBT) plus VRE (vs. WL) targeting multiple social fears was more effective at decreasing SAD severity (Bouchard et al., 2017); however, as it examined VRE blended with CBT, the efficacy of self-guided VRE as a stand-alone treatment could not be deduced.

Most of the VRE conducted so far has been therapist-led but attempts toward decreasing therapist participation have recently been underway. For example, one-session of minimally-therapist-assisted VRE (vs. *in vivo* exposure therapy) produced comparable reduction in behavioral avoidance frequency and fear severity in spider phobic individuals (Miloff et al., 2019). Likewise, completely self-guided, smartphone app-delivered VRE (vs. WL) resulted in large decreases in fear of heights for persons with acrophobia (Donker et al., 2019), with strong engagement across unique exposure scenarios that evoked varying degrees of fear (Donker, van Klaveren, Cornelisz, Kok, & van Gelder, 2020). In addition, the large effect of a relatively inexpensive, one-session, self-directed, Google Cardboard headset-delivered VRE for public speaking phobia did not differ significantly from a therapist-guided treatment (Lindner, Miloff, Fagernas, et al., 2019). Such efforts are important to test further as VR can enhance the delivery of self-help exposure therapy to persons with SAD by decreasing costs and increasing access when trained therapists are not available. For example, it could decrease costs if used as part of a stepped care model. All patients within a health management organization (HMO) who could benefit from VRE, could first be assigned to complete self-guided VRE within a clinic. The clinic might contain multiple VRE equipped rooms and overseen by one minimum wage paid staff person who serves as a technician troubleshooting any mechanical faults with the equipment. If this was not sufficient to lead to remission the patient would then be provided more intensive in-person CBT with a well-trained and qualified CBT practitioner (Boeldt, McMahon, McFaul, & Greenleaf, 2019; Newman, Szkodny, Llera, & Przeworski, 2011). Alternatively self-help VRE equipment could be rented and/or purchased and used at home as the price of VRE devices has been declining across time (Bun, Gorski, Grajewski, Wichniarek, & Zawadzki, 2017).

Accordingly, building on prior work, we aimed to examine the efficacy of self-guided VRE for persons with SAD. There are several innovative aspects of the VRE used in the current study. First, our VRE had a virtual therapist built into it who auditorily guided participants through the exposure exercises and who conveyed CBT principles (see online supplementary material (OSM) for therapist script examples). We are aware of only one prior self-help VRE controlled trial that included a virtual therapist and this was for a treatment for acrophobia (Donker et al., 2019). The use of the virtual therapist also adds to prior work on non-self-help VREs (Anderson et al., 2013; Bouchard et al., 2017; Harris et al., 2002). Relatedly, the present VRE included a built-in system to deliver and record subjective units of distress (SUDS) ratings. Further, it detected if the participant talked in response to the lifelike human character (professional actors were hired to play these roles) and the characters would wait until the participant provided verbal output before continuing the preprogrammed interaction. The current study also augments prior self-guided VRE controlled trials that did not contain cutting-edge, omnidirectional (360°), and dynamic video technology driven by machine learning optimization algorithms, biometrics, and programming rules (Lindner, Miloff, Fagernas, et al., 2019; Lister et al., 2010). Also, to our knowledge, no other self-help controlled VRE study for SAD conducted exposure to bidirectional social interactions (Lindner, Miloff, Fagernas, et al., 2019; Lister et al., 2010). Moreover, our study provided between-session exposure homework exercises to consolidate learning; a procedure mirroring face-to-face CBT but not conducted by any other self-help

VRE study to date and omitted by more than half of the prior non-self-help studies on VRE for anxiety disorders thus far (refer to meta-analysis by Benbow & Anderson, 2019).

We hypothesized that from baseline to post-treatment, self-guided VRE would result in substantially greater decrease in primary outcomes (SAD severity and job interview anxiety) when compared to WL. Further, we hypothesized that VRE (vs. WL) would lead to larger reductions in secondary outcomes, namely trait worry and depression symptom severity. In addition, we anticipated that pre-post-treatment gains would be maintained across 3-month follow-up (3MFU) and 6-month follow-up (6MFU) periods. Lastly, we conducted exploratory analyses to examine whether there would be enhancement in presence (i.e., sense of virtual world mimicking real-life scenarios) and reduction in cybersickness across VRE sessions.

1. Method

1.1. Participants

Participants were individuals with SAD[1] ($n = 44$) selected from the community and local colleges through the Subject Pool, as well as a research advertisement portal (StudyFinder) and social media platforms (Facebook, Nextdoor, Instagram, Reddit, Craigslist). Average age was 23.30 years ($SD = 9.32$, range = 18–53). Of these, 77.3% were female, and 52.3% identified as White, 22.8% as African American, and the remaining 25% as Hispanic, Latino, Asian, American Indian, or Alaska Native. Half (or 50%) of the participants had a high school diploma, 31.2% had some formal high school education, and 18.2% had a Bachelor's degree or some college education. They were compensated up to $40 worth of Amazon gift cards (pro-rated based on degree of participation) for their time and effort in the study.

1.2. Pre-Treatment clinical interview

Psychiatric Diagnoses.—The face-to-face, brief, and precise Mini International Neuropsychiatric Interview (MINI; Sheehan et al., 1998) structured psychiatric diagnostic interview was used to determine diagnoses of SAD and other mental disorders. The MINI is based on the Diagnostic and Statistical Manual–Fifth Edition (DSM-5; American Psychiatric Association, 2013) and has been validated against other psychometrically sound diagnostic interviews (First, Spitzer, Gibbon, & Williams, 1997; Robins et al., 1988). Sheehan et al. (1997) found strong sensitivity (0.81) and specificity (0.86) for a current SAD diagnosis as well as acceptable sensitivity (0.67–0.96) and strong specificity (0.86–0.99) for other disorders.

1.3. Outcome measures at pre-, post-Treatment, 3MFU, and 6MFU

SAD Symptom Severity.—The 25-item Diagnostic and Statistical Manual-Fourth Edition–Text Revised (DSM-IV-TR)-aligned Social Phobia Diagnostic Questionnaire (SPDQ; Newman, Kachin, Zuellig, Constantino, & Cashman-McGrath, 2003) was used to assess degree of fear and avoidance related to diverse social situations. It has good 2-week

---

[1]Data on diagnostic comorbidities were included in Table S1 of the OSM.

retest reliability (κ = .63), strong internal consistency (Cronbach's αs = .95, .97, .95, .96, in the present study), as well as good convergent validity and discriminant validity (Newman et al., 2003). Further, it has excellent sensitivity (82%) and specificity (85%) (Newman et al., 2003). The 20-item Social Interaction Anxiety Scale (SIAS; Mattick & Clarke, 1998) was also used to assess SAD symptom severity. Participants rated on a 5-point Likert scale (0 = *not at all characteristic or true of me* to 4 = *extremely characteristic or true of me*) their degree of anxiety regarding social interactions (e.g., "When mixing socially I am uncomfortable"). The SIAS had excellent 4-to-12-week retest reliability (r = .92) (Mattick & Clarke, 1998), good internal consistency (αs = .82, .81, .85, .83 in the present study), as well as strong convergent and discriminant validity across diverse samples (Heimberg, Mueller, Holt, Hope, & Liebowitz, 1992).

Job Interview Anxiety Severity.—The 30-item Measure of Anxiety in Selection Interviews (MASI; McCarthy & Goffin, 2004) was used to assess job interview anxiety severity. Respondents endorsed on a 5-point Likert scale (1 = *strongly disagree* to 5 = *strongly agree*) the degree to which they experienced anxiety surrounding appearance, behavioral, communication, performance, and social aspects of interviewing (e.g., "I become very uptight about having to socially interact with a job interviewer"). Higher scores denoted greater job interview anxiety. The MASI has shown acceptable internal consistency (αs = .80, .80, .84, .82 in the present study), and strong convergent and discriminant validity (McCarthy & Goffin, 2004).

Worry Symptom Severity.—The 16-item Penn State Worry Questionnaire (PSWQ; Meyer, Miller, Metzger, & Borkovec, 1990) was utilized to measure trait worry. Participants rated the extent to which they tended to worry excessively (e.g., "My worries overwhelm me") on a 5-point Likert scale (1 = *not typical at all* to 5 = *very typical*). The PSWQ has shown good internal consistency (αs = .81, .79, .86, .83 in the current study), retest reliability (r = .74–.94), convergent and discriminant validity (Hopko et al., 2003).

Depression Symptom Severity.—The 9-item Patient Health Questionnaire (PHQ-9; Kroenke, Spitzer, & Williams, 2001) was used to assess extent of depression symptom severity. Respondents rated the frequency with which they encountered depression symptoms (e.g., "Little interest or pleasure in doing things") in the past 2 weeks on a 4-point Likert scale (0 = *Not at all* to 3 = *Nearly every day*). The PHQ-9 demonstrated excellent internal consistency (αs = .80, .71, .85, .80 herein), good retest reliability (r = .84), and strong construct validity (Kroenke et al., 2001; Kroenke & Spitzer, 2002).

Qualitative Feedback.—At post-treatment, we solicited open-ended feedback on the acceptability of the VRE via two items: (a) Would you recommend this program to others who might have problems similar to yours? (Response format: Yes/No) (b) Why? (Open-ended).

1.4.   Session-by-session process measures

The following measures were administered during every VRE session.

Presence.—The 14-item iGroup Presence Questionnaire (IPQ; Schubert, Friedmann, & Regenbrecht, 2001) was administered. Respondents endorsed the extent to which they experienced the VRE scenarios to be realistic after immersing in the virtual environment for 25–30 min during each session on a 7-point Likert scale. The IPQ measured three elements of presence: *involvement* (4-item; e.g., "Somehow I felt that the virtual world surrounded me."); *realness* (4-item; e.g., "How real did the virtual world seem to you?"); *spatial presence* (6-item; e.g., "I felt present in the virtual space."). The IPQ evidenced high internal consistency (average α = .90 in this study) and construct validity (Vasconcelos-Raposo et al., 2016).

Cybersickness.—The 16-item Simulator Sickness Questionnaire (SSQ; Kennedy, Lane, Berbaum, & Lilienthal, 1993) was used. It comprises 3 domains: *disorientation* (e.g., vertigo, dizziness), *nausea* (e.g., burping, elevated salivation, nausea, stomach tightness), and *oculomotor effects* (e.g., blurred vision, eyestrain, headache, problems concentrating). Participants rated the degree to which they experienced the symptoms currently on a 4-point Likert scale (1 = *none* to 4 = *severe*). The SSQ has shown strong internal consistency (α = .98 herein) and construct validity (Bouchard, Robillard, & Renaud, 2007).

Subjective Units of Distress (SUDS).—Participants rated their level of SUDS on an 11-point Likert scale (0 = *not at all* to 10 = *extremely*) at various time-points before and during each exposure (Wolpe & Lazarus, 1967).

1.5.   Procedure

To be eligible, participants had to be 18 years old or above, provide informed consent, and meet DSM-5 criteria (American Psychiatric Association, 2013) for SAD. The SAD diagnosis was determined based on responses to the SPDQ (Newman et al., 2003) self-report at the screening stage and the face-to-face MINI (Sheehan et al., 1998) structured psychiatric diagnostic interview during the first laboratory visit. The face-to-face components of the current pilot RCT took place from December 2018 to March 2020 before the COVID-19 restrictions began to take effect. We were originally aiming for 30 participants each in the VRE and WL groups but had to stop recruiting for the study due to the COVID-19 pandemic. In addition, all three of the VRE dropouts during treatment were due to our inability to provide a full course of the VRE treatment due to COVID-19. Exclusion criteria included presence of psychosis, mania, alcohol or substance use disorder, suicidality, or any medical or organic disorder that hindered meaningful participation. In the present study, the inter-rater agreement for the SAD diagnosis by two highly trained independent raters was good (Cohen's κ = .80, 95% CI [.70,.95]). Also, inter-rater agreement herein was satisfactory-to-excellent across all other MINI-derived DSM-5-based mental disorders (average κs = .59–1.00).

After the baseline MINI interview, eligible participants completed symptom self-reports (counter-balanced to rule out order effects). Each participant was subsequently randomized to VRE or WL based on a Microsoft Excel randomizer that generated a random set of numbers used to determine treatment allocation for each participant ('1' coded for VRE and '0' for WL); an established method in clinical trials (Padhye, Cron, Gusick, Hamlin, &

Hanneman, 2009). Experimenters were blinded to the treatment allocation. Whereas eligible VRE participants began treatment immediately during the first lab visit, WL participants did not begin treatment until after four weeks (see Fig. 1 for CONSORT flow chart).

1.6. VRE for general social skills Training or job interview anxiety

We collaborated with a VR company, BehaVR (https://www.behavr.com/), that develops and provides VR-based mental health treatments (Paul, Bullock, & Bailenson, 2020; Summers, Schwartzberg, & Wilhelm, 2021). Specifically, we worked with Limbix to script the VR actors' verbal and non-verbal behaviors, virtual therapist's instructions, task prompts, and treatment rationale content in all scenes shown on a Pico Goblin VR headset. Two exposure therapy themes were created (informal dinner party or formal job interview) based on CBT principles and the literature. The headset displayed a 5.5-inch diagonal screen size (depth: 139.7 mm; height: 122 mm; width: 68 mm) with 2560×1440-pixel resolution, 3 degrees of freedom, 92° field of view, a refresh rate of 70Hz, and 54–71 mm interpupillary distance (Kyoto, 2017). The VRE also contained cutting-edge 360° stereoscopic views allowing for an omnidirectional range of visual perspective as participants could look in all directions within the virtual environment. It also contained dynamic video technology driven by machine learning optimization algorithms, biometrics, and programming rules (Lindner, Miloff, Fagernas, et al., 2019; Lister et al., 2010). With respect to biometrics, it had the ability to detect voice output (e.g., a person not speaking served as a precondition to virtual therapist prompts for participants to speak with the avatar). Also, the VRE could identify head position and movement (i.e., detected when participants positioned their heads to point to and highlight a specific SUDS Likert Scale rating, and detected when participants looked up, down, or sideways within corresponding 360-degree views). Displays of the pre-recorded VRE videos could be smoothly operated wirelessly with a tablet, monitored by the researchers, and it showed the selected scenes in real-time. Further, the headset could be conveniently switched on and off, and tracked. The exposure scenes each lasted between 1.5 and 10 min.

Participants were allowed to choose between the informal dinner party or the formal job interview themes for VRE. For both themes, each scene was developed to be more anxiety-provoking as the VRE progressed. In greater anxiety-inducing scenes occurring later in the program, the virtual interviewers and actors were directed to display less compassion, friendliness, humor, and pleasant verbal and non-verbal behaviors and demeanors to elicit elevated anxiety (Carless & Imber, 2007). Also, the voice of a virtual therapist was embedded within the VRE to coach the participants through each distinct scene by orienting and prompting them to the exposure therapy task(s), as well as conveying core principles of exposure therapy, and repeating the instructions if the participant was not responsive within 5 s. Each scene started with a paused video, during which participants were oriented by the virtual therapist to the context (see OSM for examples of therapist dialogue). The VRE also automatically administered the SUDS for participants to rate their distress before each scenario began, immediately upon first exposure, and toward the end of the exposure. At these time points, when the VRE presented the 11-point SUDS, the scene paused (i. e., characters froze) momentarily, participants positioned their head to face their chosen number on the response scale which was then highlighted, and they then pressed a button on the

VRE headset for their responses to be logged. Using the same process, participants were automatically prompted by the virtual therapist to retrospectively rate their peak, and lowest SUDS immediately after each exposure scene ended. On the whole, the VRE automatically prompted SUDs ratings before, immediately upon first exposure, peak, lowest, and toward end of exposure which were recorded within the VRE.

The dinner party exposure therapy theme comprised six scenes that encompassed four major social fear domains (Klinger et al., 2005): assertiveness fear (e.g., vocalizing a contrarian perspective), intimacy fear (e.g., making small talk), observation anxiety (e.g., performing a task under observation by others), and performance anxiety (e.g., feeling the pressure to impress others). Participants had to introduce themselves to the host of the dinner party (Scene 1; refer to Fig. 2a), order a drink in front of others without the presence of the party guide (Scene 2), ask the host for directions to the restroom (Scene 3), express dissenting opinions (Scene 4), flag the host to send a drink back (Scene 5), and speak to a person of authority (Scene 6). Table S2 and Figs. S1a–1e in the OSM detail the dinner party exposure therapy-themed scenes.

The job interview exposure therapy theme consisted of seven scenes developed based on empirically-supported exposure therapy and related CBT protocols for SAD (Hope, Heimberg, & Turk, 2019; Mayo-Wilson et al., 2014). Participants underwent these exposure therapy scenes in the following order: wait to be called for an interview with two other visibly anxious interviewees in the waiting room (Scene 1); sit quietly and by themselves in the conference room for a few minutes before the first interview began (Scene 2); answer standard questions about themselves, as well as their skill sets, potential contributions, greatest professional strength, and communication skills, in a one-on-one interview format at a conference room with a friendly interviewer (Scene 3); respond to three non-traditional questions to assess their degree of creative thinking, in a one-on-one interview format at a casual setting of the office cafeteria (Scene 4); experience a hostile and critical interviewer who first accused them of being late, and then asked two difficult questions to assess their weaknesses and degree of creative thinking, in a one-on-one interview format at a conference room (Scene 5); answer two questions on how they thought prior colleagues would appraise them and their workplace conflict resolution skills, in a group interview format with two other interviewees in a conference room (Scene 6; refer to Fig. 2b); prepare an impromptu speech for 1.5 min and deliver it to two interviewers who appeared disengaged and unfriendly during the 3-min speech delivery (Scene 7) (refer to Table S3 and Figs. S2a–S2f for more details).

1.7. Session protocol

Throughout the treatment study, no in-training, licensed, or trained human therapist was available for participants to access and the research assistant (RA) with whom the participant interacted only served as an administrator. As mentioned, all aspects of the intervention (e.g., instructions, therapy rationale, SUDS ratings) were standardized and delivered by scripted audio instructions or automated via the VRE equipment through the virtual therapist with no treatment-related human therapist needed. For instance, no human help was provided with regard to using the VRE, and the RAs were neither trained clinically

to perform cognitive behavioral therapy nor to articulate exposure therapy principles, rationales, and approaches.

VRE participants attended the sessions held in a clinical psychology lab twice a week, for 50–60 min each. During each session, participants were first greeted by an RA whose duties were to oversee completion of the research-related self-reports before and after each session and to ensure that standard audio scripts were administered at the appropriate times. The RA also advanced the VRE to the next exposure therapy scenario only after the participant showed adequate habituation using a standard algorithm (across three consecutive trials, SUDS reduced by at least 50% or SUDS before exposure was ≤ 3). At the start of each session, participants received psychoeducation about exposure therapy by listening to an audiotape and reading handouts that informed them what exposure therapy was and how it worked to treat SAD (e.g., reversing the vicious cycle of avoidance, directing participants to purposefully enter social situations that create anxiety and to stay and engage fully in those contexts). For instance, participants were instructed to focus fully on connecting with their emotions and people in the virtual environment as they would in real-life, and to allow and tolerate whatever discomfort came up during each social exposure exercise. The audiotapes emphasized in various ways the importance of *avoiding avoidance* and not doing anything to get rid of anxious feelings during each exposure session (e. g., distracting oneself, not paying close attention). As the audiotape played, participants referred to a handout illustrating the vicious cycle of avoidance (Anxiety of social situations → Person tells oneself they cannot cope with situation → Avoidance of situation provides temporary relief → Decreased confidence in being able to cope → Long-term avoidance of social opportunities) and how VRE will reverse the cycle of avoidance (Exposure to social situations → Learning to tolerate anxiety → Confidence in being able to cope → Reduced avoidance in the short- and long-term → Greater opportunities for social engagement).

In the first session, the audio directed participants to anchor their SUDS levels by recalling experiences where they would rate their SUDS to be 0 = *not at all*, 10 = *extremely*, and 5 = *moderately*, in that order. During the next 20–25 min, they wore a Pico Goblin VR headset to immerse in the VRE. The time spent in the VRE was standardly capped at 25–30 min to prevent cybersickness (Bouchard, Robillard, Loranger, & Larouche, 2012). The VRE prompted participants to rate their SUDS before, upon first exposure, immediately before ending, as well as to report their peak (highest), and lowest SUDS retrospectively after each exposure scene. If, the initial SUDS before exposure was ≤ 3 across three consecutive trials, the participant was advanced to the next exposure scenario. If any initial SUDS before exposure was ≥ 4, the same exposure scenario was repeated until habituation (defined as at least 50% reduction from peak SUDS across 3 consecutive trials) was achieved. Thus, participants engaged in at least three trials for each exposure scene. At the end of each lab visit, participants completed the process measures of presence, cybersickness, self-focused attention, other-focused attention, self-efficacy, and feared negative prediction. WL participants began Session 1 about four weeks post-randomization. In total, VRE participants engaged in 1–8 exposure therapy sessions ($M = 3.55$, $SD = 1.68$, median = 4.00). Treatment was mutually terminated only after participants displayed habituation following 2 cycles of the VRE and engagement with homework, described below.

1.8.   Pre-to post-Treatment between-session homework

Participants were prescribed between-session *in vivo* exposure therapy homework to consolidate within-session learning. Three minutes of standardized audio rationale and instructions for homework was provided at the end of session. This homework highlighted the value of practicing the skills participants learned from the VRE to their everyday lives between sessions. It emphasized the importance of participants putting themselves out there to feel more confident and to stay in the situation long enough to make it more likely that their anxiety would decrease (ideally decreasing 50% from their initial anxiety when entering the situation). The homework emphasized that the actual exposure exercise outcome (e.g., receiving a positive response to asking a person out for a date) did not matter, and emphasized the importance of remaining in anxiety-inducing social situations to build confidence in their ability to cope. Each participant referred to a handout and was instructed to select at least one unique exposure therapy scenario to engage with at least once before the next session. To promote feasibility of *in vivo* exposure, participants were instructed to commit to fulfilling an achievable behavioral goal, to plan and specify their real-life exposure therapy activities, as well as the place, day, and time for completion. Further, to increase compliance, participants were instructed to anticipate obstacles to homework completion and brainstorm solution (s). Participants were instructed that for each homework exposure task they were to record their SUDS and self-efficacy at various time-points of the exposure (e.g., before undergoing the behavioral task), wherein some ratings (i.e., upon first exposure, peak, lowest, toward the end) were retrospectively recalled immediately after the exposure ended.

1.9.   Research fidelity monitoring

Three Ph.D./Psy.D. clinical psychology candidates ensured high fidelity to delivery of the research protocols. Further, the doctoral candidates trained and supervised 7 RAs by role-playing the initial psychiatric diagnostic interview, and how to administer the standard protocol (e.g., algorithm for advancement to the next scene) and research-related measures during each session. Also, regular group RA meetings were conducted to address any research protocol-related problems, offer more training if needed, and to preserve standardization of research protocol delivery. All sessions were videotaped. Throughout the process, a Ph.D. candidate watched 15% of the tapes to offer corrective feedback if required to ensure that the RAs followed the standardized research protocol. Additionally, the doctoral candidates ensured that the research team requested that participants completed the 3MFU and 6MFU self-reports online.

1.10.   Data analyses

All statistical analyses were carried out using *R* version 3.6.3 (R Core Team, 2020). The *nlme R* package (Pinheiro, Bates, DebRoy, & Sarkar, 2020 & R Core Team, 2020) was used to conduct multilevel modeling analyses given repeated measures nested within persons and treatment groups (Enders, 2017). Treatment groups were coded as '0' (WL) and '1' (VRE). Time-points were coded as '0' (pre-treatment), '1' (post-treatment), '2' (3MFU), and '3' (6MFU). To ascertain treatment effect on dimensional SAD as a primary outcome, we calculated a composite SAD symptom severity score by averaging standardized SPDQ

and SIAS scores to refrain from multiple testing and prevent Type I error (Stolz et al., 2018). Other outcomes were indexed by single measures: job interview anxiety (MASI); trait worry (PSWQ); depression severity (PHQ-9). Regarding time trends, the multilevel models included the intercept, time, group, and time by group interaction as fixed effects, and intercept as a random effect. We also determined if the proportion of participants with clinically significant change (Jacobson & Truax, 1991) differed between groups from pre-to-post-treatment. Missing data (17% in total) due to dropouts and COVID-19 restrictions (refer to CONSORT flowchart in Fig. 1) was managed with multiple imputation via a predictive mean matching algorithm based on Rubin's (1987) rules using the *mice R* package (van Buuren & Groothuis-Oudshoorn, 2011). Multilevel multiple (vs. single) imputation minimizes bias (e.g., reduces standard error) and enhances power and efficiency in estimating random and fixed effects of continuous outcomes (Audigier et al., 2018), especially for treatment studies with four or more time-points comprising $<30\%$ of missing data (Collins, Schafer, & Kam, 2001; Enders, 2011; Gottfredson, Sterba, & Jackson, 2017). To this end, multilevel multiple imputation yields more credible estimates. As recommended, 50 imputations, each with 100 iterations based on the Markov chain Monte Carlo method, were used (Grund, Lüdtke, & Robitzsch, 2018). To ease interpretation, Hedge's *g* was computed, such that values of 0.2, 0.5, and 0.8 denoted small, moderate, and large effect sizes, respectively (Dunst, Hamby, & Trivette, 2004).

1.11. Power analysis

Following recommendations (Arend & Schafer, 2019), an *a priori* power analysis was conducted that mirrored study conditions (e.g., attrition rate of 20–30%, sample size of $N = 44$) with the *R* package *lmmpower* (Magnusson, 2018). After 5000 replications per condition, we observed 85.12–92.62% power to identify significant Group $\times$ Time interaction effects with $g = 0.5$. Thus, the current pilot RCT was well-powered.

2. Results

2.1. Descriptive data and initial group differences

Fig. 1 presents the CONSORT flowchart of this pilot RCT. In total, the attrition rate was 20%.[2] Table 1 displays descriptive data for each measure in each group across time. No initial group differences were observed in the SAD composite ($\beta(SE) = 0.43$ (0.22), $p = .054$, $g = 0.60$), MASI ($\beta(SE) = 4.41$ (4.90), $p = .370$, $g = 0.28$), and PHQ-9 ($\beta(SE) = 2.00$ (2.24), $p = .27$, $g = 0.34$). However, pre-treatment PSWQ was significantly higher for VRE than WL ($\beta(SE) = 7.41$ (1.81), $p = .002$, $g = 1.00$).

3. Primary outcomes

Table 2 as well as Figs. S3–S6 present the simple slopes of multilevel model results for all time-points. Table 3 displays the multilevel model pre-to-post-treatment results.

---

[2]Due to space constraints, attrition analyses were included in Appendix A of the OSM.

SAD Severity Composite.

During pre-to-post-treatment, a significant Group × Time interaction effect emerged (β(*SE*) = −0.97 (0.28), *p* < .001, *g* = −1.06). Simple slope analyses indicated significantly larger VRE decrease in SAD severity composite score (β(*SE*) = −1.84 (0.17), *p* < .001, *g* = −4.77) relative to WL (β(*SE*) = −0.86 (0.21), *p* = .002, *g* = −2.35). Further, showing maintenance of gains, the within-VRE SAD severity composite score did not significantly change during post-3MFU (β(*SE*) = 0.01 (0.05), *p* = .765, *g* = 0.12) and post-6MFU (β(*SE*) = −0.04 (0.14), *p* = .747, *g* = −0.13). VRE (vs. WL) had a marginally significantly higher proportion of participants with clinically reliable improvement on the SAD composite severity (Reliable recovery: 88.46% vs. 66.67%; No change: 11.54% vs. 33.33%; $\chi^2(df = 1) = 3.11$, *p* = .078, *g* = 0.54).

Job Interview Anxiety.

From pre-to-post-treatment, a significant Group × Time interaction effect emerged (β(*SE*) = −13.95 (6.48), *p* = .037, *g* = −0.66). Simple slope analyses indicated that decrement in MASI was larger in VRE (β(*SE*) = −36.28 (3.83), *p* < .001, *g* = −4.17) relative to WL (β(*SE*) = −21.59 (4.82), *p* < .001, *g* = −2.51). Also, within the VRE group, MASI substantially decreased at pre-3MFU (β(*SE*) = −16.33 (1.94), *p* < .001, *g* = −3.70) and pre-6MFU (β(*SE*) = −12.35 (1.44), *p* < .001, *g* = −3.77). Showing maintenance of gains, there were no significant changes in MASI at post-3MFU (β(*SE*) = − 0.28 (1.03), *p* = .791, *g* = − 0.10) and post-6MFU (β(*SE*) = −4.41 (3.22), *p* = .184, *g* = −0.53). Contrary to prediction, VRE (vs. WL) did not significantly differ in terms of proportion of participants with clinically reliable improvement on the MASI (Reliable recovery: 84.62% vs. 66.67%; No change: 15.38% vs. 33.33%; Reliable deterioration: 0.00% vs. 0.00%; χ^2 (df = 1) = 1.95, *p* = .163, *g* = 0.42).

3.1. Secondary outcomes

Trait worry.—From pre-to-post-treatment, a significant Group × Time interaction effect was found (β(*SE*) = −8.43 (3.33), *p* = .015, *g* = −0.78). Simple slope analyses showed that reduction in PSWQ was notably greater in VRE (β(*SE*) = −21.28 (2.11), *p* < .001, *g* = −4.45) than WL (β(*SE*) = −13.17 (2.64), *p* < .001, *g* = −2.80). Also, PSWQ significantly reduced at pre-3MFU (β(*SE*) = −9.24 (1.24), *p* < .001, *g* = −3.27) and pre-6MFU (β(*SE*) = −8.19 (1.10), *p* < .001, *g* = −3.28). Showing maintenance of gains, no significant changes were observed from post-3MFU (β(*SE*) = −1.71 (1.02), *p* = .107, *g* = −0.65), and post-6MFU (β(*SE*) = −6.00 (3.51), *p* = .100, *g* = −0.66). The proportion of participants who had clinically reliable improvement in PSWQ scores did not significantly differ between VRE and WL from pre-to-post-treatment (Reliable recovery: 88.46% vs. 72.22%; No change: 11.54% vs. 27.78%; Reliable deterioration: 0.00% vs. 0.00%; $\chi^2(df = 1) = 1.89$, *p* = .170, *g* = 0.42).

Depression severity.—From pre-to-post-treatment, no significant Group × Time interaction effect was observed (β(*SE*) = −3.61 (2.05), *p* = .087, *g* = −0.54). Simple slope analyses indicated that PHQ-9 decrement was significant in VRE (β(*SE*) = −6.70 (1.34), *p* < .001, *g* = −2.21) but not WL (β(*SE*) = −2.79 (1.44), *p* = .079, *g* = −1.09). Further, within the VRE group, PHQ-9 considerably decreased at pre-3MFU (β(*SE*) = −2.88 (0.57), *p* <.001, *g*

= −2.21) and pre-6MFU (β(*SE*) = −2.28 (0.50), *p* <.001, *g* = −2.00). Showing maintenance of gains, no significant shifts in PHQ-9 were found at post-3MFU (β(*SE*) = −0.12 (0.32), *p* = .711, *g* = −0.15) and post-6MFU (β(*SE*) = −1.24 (0.95), *p* = .202, *g* = −0.51) within the VRE.[3] VRE (vs. WL) did not significantly predict clinically significant change in PHQ-9 from pre-to-post-treatment (Reliable recovery: 57.69% vs. 50.00%; No change: 42.31% vs. 38.89%; Reliable deterioration: 0.00% vs. 11.11%; $\chi^2(df = 2)$ = 3.04, *p* = .219, *g* =0.53).[4],[5]

3.2. Feasibility measures

Number of scenes.—Participants initially randomly assigned to the VRE (*n* = 26) completed a range from 3 to 55 scenes, for an average of 18 scenes (*SD* = 14). Fifteen participants (58%) reached scenes of the highest difficulty level.

Presence.—From Sessions 1 to 5, significant moderate increases across time were observed for spatial presence (β(*SE*) = 0.074 (0.031), *p* = .019, *g* = 0.36), involvement (β(*SE*) = 0.093 (0.039), *p* = .017, *g* = 0.36), and realness (β(*SE*) = 0.092 (0.031), *p* = .003, *g* = 0.45). Moreover, Table 4 shows that spatial presence (*g* = 0.56) and realness (*g* = 0.83) in the current RCT were significantly higher than an earlier original study (Regenbrecht & Schubert, 2002). However, no significant difference between studies were found for involvement (*g* = 0.37) in comparison to Regenbrecht and Schubert (2002). Also, presence facets did not significantly differ from Mostajeran, Balci, Steinicke, Kuhn, and Gallinat (2020; *g* = −0.19–0.25). Table 4 details these between-study findings.

Cybersickness.—From Sessions 1 to 5, a significant moderate decrease over time was found for cybersickness (β(*SE*) = −0.58 (0.24), *p* < .001, *g* = −0.43). Moreover, cybersickness from the current RCT was significantly lower than data from Bouchard, St-Jacques, Renaud, and Wiederhold (2009) who recruited patients with specific phobias (*g* = −0.84), but not patients with SAD in Kim et al. (2017; *g* = 0.02; see Table 4).

Homework compliance.—Eighty percent of participants (*n* = 21) randomly assigned to the VRE demonstrated homework compliance. The number of separate between-session *in vivo* exposure exercises that participants completed ranged from 1 to 13 (*M* = 3.96, *SD* = 4.27).

Qualitative Feedback.—A large proportion of participants (85%) mentioned that they would recommend the VRE to others struggling with elevated SAD (refer to Table S6 for detailed feedback).

---

[3]More information on RCI outcomes were included in Table S7 of the OSM.

[4]Multilevel models that included both random intercept and random slope of time – an approach that reduces biased and anticonservative parameter and standard error estimates (Bell, Fairbrother, & Jones, 2018) – yielded similar results, as shown in Tables S8 and S9.

[5]All findings were similar when we conducted sensitivity analyses using full information maximum likelihood (FIML) instead of multiple imputation. FIML uses all available data but does not strategically replace missing values (Graham, 2009).

4. Discussion

The present study tested the efficacy of self-guided VRE as an independent treatment for persons with SAD with diverse interpersonal fears. Findings showed that from baseline to post-treatment, VRE (vs. WL) produced greater reductions in SAD severity (indexed by SPDQ and SIAS), job interview anxiety, and trait worry, but not depression symptoms. Significant pre-post-treatment effect sizes between groups ranged from moderate-to-large dimensionally ($g = -1.11$ to -0.66), and categorically (indexed by reliable change indices; $g = 0.42$ to 0.54). Further, the non-significant between-group effect of VRE (vs. WL) on depression severity was moderate dimensionally ($g = -0.54$) and categorically ($g = 0.53$). Moreover, within-group analyses revealed that although VRE led to significant change in depression pre to post, there was no change in the wait-list group. Furthermore, VRE effected large decrements in SAD symptom severity, job interview anxiety, trait worry, and depression symptom severity from baseline to 3-month and 6-month follow-up time-points (gs $= -4.60$ to -2.00). Further, maintenance of treatment gains was observed. Overall, our results are consistent with data pooled across 7 studies demonstrating that at least three sessions of VRE were superior to WL at reducing social fear and avoidance symptoms for persons with SAD (Hedge's $g = -1.70$ to -0.76; Horigome et al., 2020; Morina, Kampmann, Emmelkamp, Barbui, & Hoppen, 2021). Correspondingly, participants' feedback largely suggested that the VRE was acceptable, presence and cybersickness levels were better or comparable to other VRE, most participants were compliant with homework, and a majority reached their most difficult scenes, demonstrating feasibility of the self-help protocol.

SAD and job interview anxiety severity likely improved during and after VRE due to the accumulation of corrective experiences (i.e., disconfirmed expectations of negative outcomes) in diverse fear-evoking social interactions (cf. *emotional processing theory*; Foa & McLean, 2016). Relatedly, participants likely learned that distress associated with social interactions was tolerable and would eventually abate if they remained in the situation and allowed themselves to experience the full range of emotions that might arise in various anxiety-inducing social situations. Indeed, prior exposure therapy for SAD trials evidenced that reduced SAD severity over time coincided with decline in threat probability overestimation and anticipatory processing (Wong et al., 2017). Further, pre-post treatment gains could be explained by the occurrence of within- and between-session habituation that has been shown to reliably predict enhanced exposure therapy outcomes (see recent meta-analysis by Rupp, Doebler, Ehring, & Vossbeck-Elsebusch, 2017). Another account might be that, irrespective of habituation, VRE altered emotion-laden memory structures such that novel, safety-based learned connections among previously-feared and avoided situations inhibited prior threat-based connections (cf. *inhibitory learning theory*; Craske, Treanor, Conway, Zbozinek, & Vervliet, 2014). Of note, habituation and/or extinction learning likely occurred via the virtual therapist repeatedly emphasizing the goal of purposefully feeling anxious during exposure exercises in the service of major and long-term aims (e.g., attaining a dream job) and the importance of tolerating anxiety. Also, SAD symptom improvement might have occurred via the audiotapes reiterating and elaborating on how exposure therapy reverses the cycle of avoidance during psychoeducation at the start of each session. Further,

SAD severity enhancements may be due to the recurrent homework assignments that detailed instructions to practice exposure exercises irrespective of the actual outcomes and to stay in the situation until their anxiety reduced.

Simultaneously, changes in specific cognitive and behavioral processes might explain why SAD and job interview anxiety severity declined more in VRE than WL. For instance, it is plausible that the VRE might have led to decrements in self-focused attention, as proposed by *cognitive theories* (Norton & Abbott, 2016), and observed in CBT for SAD (Gregory & Peters, 2017). In addition, it is possible that reduced hyper-self-monitoring during exposure therapy co-occurred with decreased focus on anxiety symptoms (Hofmann, 2000), thus leading to lower SAD symptoms from baseline to post-treatment and follow-up periods. On that note, the self-guided VRE and audiotapes highlighted how essential it was to interact with the human-like characters mindfully as they would in real-life settings and to eliminate any safety behaviors (e.g., distracting oneself or being overly self-conscious). Similarly, persons might have raised their self-efficacy during the VRE (Kampmann, Emmelkamp, & Morina, 2019), and thus experienced less fear and coped more effectively with myriad interpersonal situations. Noteworthy was the fact that the virtual therapist and audiotapes in the current study impressed upon participants how engaging in more frequent exposure exercises could help them to build and gain confidence to tolerate, enter, and manage more social situations. Future studies should incorporate process measures of pertinent constructs to test mechanisms of change in VRE for SAD and to clarify the SAD profiles for whom VRE would be helpful.

Additionally, why did VRE lead to reductions in trait worry across baseline to post-treatment and follow-up time-points? This was probably because participants randomly assigned to VRE (vs. WL) had higher trait worry at pre-treatment, such that reductions over time reflected regression to the mean. Alternatively, such results were unsurprising given heightened trait worry in SAD (Starcevic et al., 2007). Moreover, findings aligned with prior similar studies that showed decline in perceived stress, worry, or academic anxiety over the course of VRE for community-dwelling adults (Kampmann et al., 2016), adolescents, and children (Whiteside et al., 2020). Plausibly, our unguided VRE effected changes in trait worry as the program helped persons with SAD unlearn unhelpful perseverative thinking habits while enhancing present-mindedness and tolerance for uncertainty. These ideas are supported by recent successful RCTs of self-help Internet-delivered extinction therapy based on operant conditioning principles (Andersson et al., 2017, 2020). Future RCTs can test these notions.

Interestingly spatial presence and realness increased across sessions and cybersickness decreased. To our knowledge, these are novel findings in the literature. VRE RCTs and open trials conducted to date typically measured presence and cybersickness once at baseline but have not assessed if global presence or its specific facets changed throughout treatment (refer to reviews by Botella, Fernández-Álvarez, Guillén, García-Palacios, & Baños, 2017; Ling, Nefs, Morina, Heynderickx, & Brinkman, 2014). Thus, we showed that presence and cybersickness are malleable constructs amenable to change across sessions during the period participants were actively receiving VRE. The presence finding is consistent with evidence that experiencing presence in virtual environments assisted people with feeling involved as

they focused on the aims, procedures, exposure exercises, and related behavioral tasks of the VRE (Kang et al., 2021). Also, rise in presence might be partly due to the fact that technical glitches (e.g., frozen videos) were collaboratively eliminated prior to commencing the pilot trial. Additionally, increased presence might be due to the cutting-edge, real-world video processing and dynamic delivery of the BehaVR headset. Although we have no data supporting the superiority of a cloud-based Dynamic Experience Engine (DEE), it is possible that the use of DEE rules, engagement optimization algorithms, and biometrics, contributed to a sense of realness in the VRE environment. It is important for VR research to continue to aim toward smooth transitions from the real world to virtual environments. Relatedly, decline in cybersickness across treatment in the current study might imply that any habituation or extinction effects occurring during VRE can encompass reductions in both anxiety and cybersickness symptoms.

Limitations of the current study merit attention. First, the study had a relatively small sample size; nonetheless, the statistically significant moderate-to-large effect sizes implied that the current pilot RCT was well-powered, and this was also reflected in our *a priori* power analysis. Subsequent investigations would benefit from increased sample size (Page & Coxon, 2016). Second, the use of a WL control group could inflate effect sizes (Cunningham, Kypri, & McCambridge, 2013); thus, future VRE RCTs are encouraged to use an active control (Nolet, Corno, & Bouchard, 2020). The study also did not include a behavioral avoidance task (BAT; Castagna, Davis, & Lilly, 2017) to assess participants' head-on confrontations with actual social scenarios before and after treatment. Future VRE studies could thus include a BAT to offer further evidence that VRE-derived gains transfer to real-life situations. Also, although there was no trained therapist available, participants did have contact with an RA. Thus, it is possible that the treatment benefitted from human contact. At the same time, all aspects of the current study that were handled by an RA (e.g., transiting clients from one exposure therapy scenario to the next using our habituation algorithm) could be easily automated by embedding conditional rules into the VRE program and via streaming audio or video turned on and off by participants prior to and following each virtual exposure. Shortcomings notwithstanding, strengths of the current pilot study included the gold-standard RCT method, use of a psychiatric diagnostic interview, and the self-guided nature of the intervention. Additionally, the present pilot RCT had a high retention rate (i.e., attrition rate of 20% was akin to *in vivo* exposure therapy and other VREs; Benbow & Anderson, 2019) and low deterioration rate (i.e., current deterioration rate of 0–4% for anxiety measures matched prior meta-analytic data; Fernandez-Alvarez et al., 2019). Further, the current pilot RCT included a 6-month follow-up, and it had elevated presence and reduced cybersickness relative to prior studies (Bouchard et al., 2009; Regenbrecht & Schubert, 2002) that did not similarly use a stereoscopic 360° camera. Also, although the study protocol was administered in a lab, the immersive and algorithm-driven self-guided VRE with its automated instructions on all aspects of the treatment (e.g., treatment rationale, homework delivery, etc.) could be delivered at any location.

If findings were to be replicated, several clinical implications deserve consideration. Unguided, brief, self-help VRE has the potential to improve dissemination of exposure therapy, since most therapists in the community lack organizational incentive, training, and proficiency to optimally deliver exposure therapy or prefer to offer relaxation-based

treatments (Becker-Haimes et al., 2017). Provision of access to the self-directed VRE equipment could be provided by setting it up at a venue (e.g., dedicated rooms with computers, VRE equipment, and detailed instructions to navigate exposure exercises within- and between-sessions) within a primary care facility. Staff could be available to troubleshoot technical problems and maintain the site. Relatedly, self-directed VRE embedded in cellphones that are tailored to the person's chief presenting SAD-related avoidance problems could be conveniently inserted into VRE head mounted devices utilized in those facilities. Choosing self-directed VRE as the first course of action for SAD in a stepped-care primary treatment facility may ultimately offer greater access to diverse communities at lower cost by requiring zero or fewer face-to-face CBT sessions (Boeldt et al., 2019; Fairburn & Patel, 2017; Lindner et al., 2017). Alternatively, the equipment and access to the audio content, could be rented to participants for a delimited period of time and used within their homes. Another possibility is for CBT practitioners to incorporate the VRE program into their clinical care as an adjunctive treatment to allow for repeatable social exposure without adversely affecting therapeutic alliance (Maples-Keller, Bunnell, Kim, & Rothbaum, 2017).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (DSM-5). Washington, DC: American Psychiatric Association. 10.1176/appi.books.9780890425596

Anderson PL, Price M, Edwards SM, Obasaju MA, Schmertz SK, Zimand E, et al. (2013). Virtual reality exposure therapy for social anxiety disorder: A randomized controlled trial. Journal of Consulting and Clinical Psychology, 81, 751–760. 10.1037/a0033559 [PubMed: 23796315]

Anderson P, Rothbaum BO, & Hodges LF (2003). Virtual reality exposure in the treatment of social anxiety. Cognitive and Behavioral Practice, 10, 240–247. 10.1016/S1077-7229(03)80036-6

Andersson E, Hedman E, Wadström O, Boberg J, Andersson EY, Axelsson E, et al. (2017). Internet-based extinction therapy for worry: A randomized controlled trial. Behavior Therapy, 48, 391–402. 10.1016/j.beth.2016.07.003 [PubMed: 28390501]

Andersson E, Ljótsson B, Hedman-Lagerlöf M, Nygren L, Persson M, Rosengren K, et al. (2020). Targeting excessive worry with internet-based extinction therapy: A randomised controlled trial with mediation analysis and economical evaluation. Psychological Medicine, 1–11. 10.1017/S0033291720000781

Arend MG, & Schafer T (2019). Statistical power in two-level models: A tutorial based on Monte Carlo simulation. Psychological Methods, 24, 1–19. 10.1037/met0000195 [PubMed: 30265048]

Audigier V, White IR, Jolani S, Debray TPA, Quartagno M, Carpenter J, et al. (2018). Multiple imputation for multilevel data with continuous and binary variables. Statistical Science, 33, 160–183. 10.1214/18-sts646

Becker-Haimes EM, Okamura KH, Wolk CB, Rubin R, Evans AC, & Beidas RS (2017). Predictors of clinician use of exposure therapy in community mental health settings. Journal of Anxiety Disorders, 49, 88–94. 10.1016/j.janxdis.2017.04.002 [PubMed: 28475946]

Bell A, Fairbrother M, & Jones K (2018). Fixed and random effects models: Making an informed choice. Quality and Quantity, 53, 1051–1074. 10.1007/s11135-018-0802-x

Benbow AA, & Anderson PL (2019). A meta-analytic examination of attrition in virtual reality exposure therapy for anxiety disorders. Journal of Anxiety Disorders, 61, 18–26. 10.1016/j.janxdis.2018.06.006 [PubMed: 30646997]

Boeldt D, McMahon E, McFaul M, & Greenleaf W (2019). Using virtual reality exposure therapy to enhance treatment of anxiety disorders: Identifying areas of clinical adoption and potential obstacles. Frontiers in Psychiatry, 10, 773. 10.3389/fpsyt.2019.00773 [PubMed: 31708821]

Botella C, Fernández-Álvarez J, Guillén V, García-Palacios A, & Baños R (2017). Recent progress in virtual reality exposure therapy for phobias: A systematic review. Current Psychiatry Reports, 19, 42. 10.1007/s11920-017-0788-4 [PubMed: 28540594]

Bouchard S, Dumoulin S, Robillard G, Guitard T, Klinger E, Forget H, et al. (2017). Virtual reality compared with in vivo exposure in the treatment of social anxiety disorder: A three-arm randomised controlled trial. British Journal of Psychiatry, 210, 276–283. 10.1192/bjp.bp.116.184234

Bouchard S, Robillard G, Loranger C, & Larouche S (2012). Description of a treatment manual for in virtuo exposure with specific phobia. In Eichenberg C (Ed.), Virtual reality in psychological, medical and pedagogical applications (pp. 81–108). Budapest, Hungary: IntechOpen. 10.5772/46417.

Bouchard S, Robillard G, & Renaud P (2007). Revising the factor structure of the simulator sickness questionnaire. Annual Review of CyberTherapy and Telemedicine, 5, 128–137.

Bouchard S, St-Jacques J, Renaud P, & Wiederhold BK (2009). Side effects of immersions in virtual reality for people suffering from anxiety disorders. Journal of CyberTherapy and Rehabilitation, 2, 127–137.

Bun P, Gorski F, Grajewski D, Wichniarek R, & Zawadzki P (2017). Low-cost devices used in virtual reality exposure therapy. Procedia Computer Science, 104, 445–451. 10.1016/j.procs.2017.01.158

van Buuren S, & Groothuis-Oudshoorn K (2011). Mice: Multivariate imputation by chained equations in R. Journal of Statistical Software, 45, 1548–7660. 10.18637/jss.v045.i03

Carless SA, & Imber A (2007). The influence of perceived interviewer and job and organizational characteristics on applicant attraction and job choice intentions: The role of applicant anxiety. International Journal of Selection and Assessment, 15, 359–371. 10.1111/j.1468-2389.2007.00395.x

Castagna PJ, Davis TE, & Lilly ME (2017). The behavioral avoidance task with anxious youth: A review of procedures, properties, and criticisms. Clinical Child and Family Psychology Review, 20, 162–184. 10.1007/s10567-016-0220-3 [PubMed: 27995381]

Collins LM, Schafer JL, & Kam C-M (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. Psychological Methods, 6, 330–351. 10.1037/1082-989X.6.4.330 [PubMed: 11778676]

Craske MG, Kircanski K, Zelikowsky M, Mystkowski J, Chowdhury N, & Baker A (2008). Optimizing inhibitory learning during exposure therapy. Behaviour Research and Therapy, 46, 5–27. 10.1016/j.brat.2007.10.003 [PubMed: 18005936]

Craske MG, Treanor M, Conway CC, Zbozinek T, & Vervliet B (2014). Maximizing exposure therapy: An inhibitory learning approach. Behaviour Research and Therapy, 58, 10–23. 10.1016/j.brat.2014.04.006 [PubMed: 24864005]

Cunningham JA, Kypri K, & McCambridge J (2013). Exploratory randomized controlled trial evaluating the impact of a waiting list control design. BMC Medical Research Methodology, 13, 150. 10.1186/1471-2288-13-150 [PubMed: 24314204]

Donker T, Cornelisz I, van Klaveren C, van Straten A, Carlbring P, Cuijpers P, et al. (2019). Effectiveness of self-guided app-based virtual reality cognitive behavior therapy for acrophobia: A randomized clinical trial. JAMA Psychiatry, 76, 682–690. 10.1001/jamapsychiatry.2019.0219 [PubMed: 30892564]
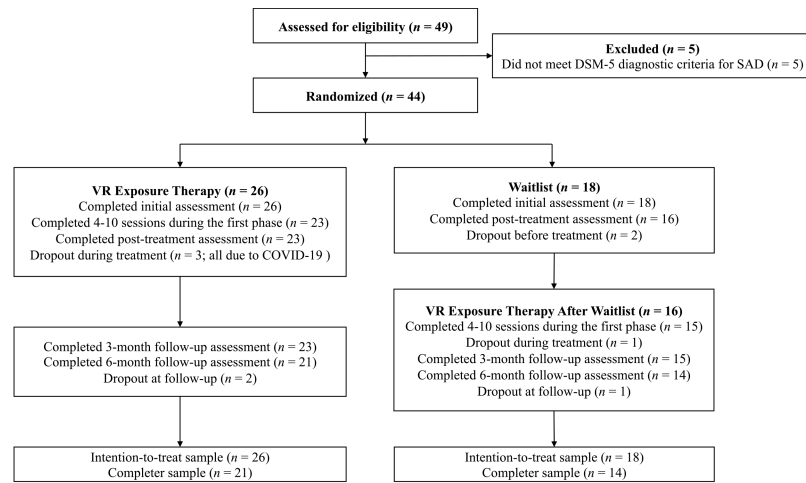
Donker T, van Klaveren C, Cornelisz I, Kok RN, & van Gelder J-L (2020). Analysis of usage data from a self-guided app-based virtual reality cognitive behavior therapy for acrophobia: A randomized controlled trial. Journal of Clinical Medicine, 9, 1614. 10.3390/jcm9061614

Dunst CJ, Hamby DW, & Trivette CM (2004). Guidelines for calculating effect sizes for practice-based research syntheses. Centerscope, 3, 1–10.

Enders CK (2011). Missing not at random models for latent growth curve analyses. Psychological Methods, 16, 1–16. 10.1037/a0022640 [PubMed: 21381816]

Enders CK (2017). Multiple imputation as a flexible tool for missing data handling in clinical research. Behaviour Research and Therapy, 98, 4–18. 10.1016/j.brat.2016.11.008 [PubMed: 27890222]

Fairburn CG, & Patel V (2017). The impact of digital technology on psychological treatments and their dissemination. Behaviour Research and Therapy, 88, 19–25. 10.1016/j.brat.2016.08.012 [PubMed: 28110672]

Fernandez-Alvarez J, Rozental A, Carlbring P, Colombo D, Riva G, Anderson PL, et al. (2019). Deterioration rates in virtual reality therapy: An individual patient data level meta-analysis. Journal of Anxiety Disorders, 61, 3–17. 10.1016/j.janxdis.2018.06.005 [PubMed: 30057346]

First MB, Spitzer RL, Gibbon M, & Williams JBW (1997). Structured clinical interview for DSM-IV axis I disorders. Washington, DC: American Psychiatric Press.

Foa EB, & Kozak MJ (1986). Emotional processing of fear: Exposure to corrective information. Psychological Bulletin, 99, 20–35. 10.1037/0033-2909.99.1.20 [PubMed: 2871574]

Foa EB, & McLean CP (2016). The efficacy of exposure therapy for anxiety-related disorders and its underlying mechanisms: The case of OCD and PTSD. Annual Review of Clinical Psychology, 12, 1–28. 10.1146/annurev-clinpsy-021815-093533

Furlong TM, Richardson R, & McNally GP (2016). Habituation and extinction of fear recruit overlapping forebrain structures. Neurobiology of Learning and Memory, 128, 7–16. 10.1016/j.nlm.2015.11.013 [PubMed: 26690954]

Garcia-Palacios A, Botella C, Hoffman H, & Fabregat S (2007). Comparing acceptance and refusal rates of virtual reality exposure vs. in vivo exposure by patients with specific phobias. CyberPsychology and Behavior, 10, 722–724. 10.1089/cpb.2007.9962 [PubMed: 17927544]

Gottfredson NC, Sterba SK, & Jackson KM (2017). Explicating the conditions under which multilevel multiple imputation mitigates bias resulting from random coefficient-dependent missing longitudinal data. Prevention Science, 18, 12–19. 10.1007/s11121-016-0735-3 [PubMed: 27866307]

Graham JW (2009). Missing data analysis: Making it work in the real world. Annual Review of Psychology, 60, 549–576. 10.1146/annurev.psych.58.110405.085530

Gregory B, & Peters L (2017). Changes in the self during cognitive behavioural therapy for social anxiety disorder: A systematic review. Clinical Psychology Review, 52, 1–18. 10.1016/j.cpr.2016.11.008 [PubMed: 27912159]

Grund S, Lüdtke O, & Robitzsch A (2018). Multiple imputation of missing data for multilevel models: Simulations and recommendations. Organizational Research Methods, 21, 111–149. 10.1177/1094428117703686

Harris SR, Kemmerling RL, & North MM (2002). Brief virtual reality therapy for public speaking anxiety. CyberPsychology and Behavior, 5, 543–550. 10.1089/109493102321018187 [PubMed: 12556117]

Heimberg RG, Mueller GP, Holt CS, Hope DA, & Liebowitz MR (1992). Assessment of anxiety in social interaction and being observed by others: The social interaction anxiety scale and the social phobia scale. Behavior Therapy, 23, 53–73. 10.1016/S0005-7894(05)80308-9

Hembree EA, & Cahill SP (2007). Chapter 17 - obstacles to successful implementation of exposure therapy. In Richard DCS, & Lauterbach D (Eds.), Handbook of exposure therapies (pp. 389–408). Burlington: Academic Press. 10.1016/B978-012587421-2/50018-1.

Hofmann SG (2000). Self-focused attention before and after treatment of social phobia. Behaviour Research and Therapy, 38, 717–725. 10.1016/S0005-7967(99)00105-9 [PubMed: 10875193]

Hope DA, Heimberg RG, & Turk CL (2019). Managing social anxiety: A cognitive- behavioral therapy approach (3rd ed.). New York, NY: Oxford University Press.

Hopko DR, Reas DL, Beck JG, Stanley MA, Wetherell JL, Novy DM, et al. (2003). Assessing worry in older adults: Confirmatory factor analysis of the Penn State Worry Questionnaire and psychometric properties of an abbreviated model. Psychological Assessment, 15, 173–183. 10.1037/1040-3590.15.2.173 [PubMed: 12847777]

Horigome T, Kurokawa S, Sawada K, Kudo S, Shiga K, Mimura M, et al. (2020). Virtual reality exposure therapy for social anxiety disorder: A systematic review and meta-analysis. Psychological Medicine, 50, 1–11. 10.1017/S0033291720003785 [PubMed: 31663847]

Jacobson NC, & Newman MG (2017). Anxiety and depression as bidirectional risk factors for one another: A meta-analysis of longitudinal studies. Psychological Bulletin, 143, 1155–1200. 10.1037/bul0000111 [PubMed: 28805400]

Jacobson NS, & Truax P (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. Journal of Consulting and Clinical Psychology, 59, 12–19. 10.1037//0022-006X.59.1.12 [PubMed: 2002127]

Kampmann IL, Emmelkamp PM, Hartanto D, Brinkman WP, Zijlstra BJ, & Morina N (2016). Exposure to virtual social interactions in the treatment of social anxiety disorder: A randomized controlled trial. Behaviour Research and Therapy, 77, 147–156. 10.1016/j.brat.2015.12.016 [PubMed: 26752328]

Kampmann IL, Emmelkamp PMG, & Morina N (2019). Cognitive predictors of treatment outcome for exposure therapy: Do changes in self-efficacy, self-focused attention, and estimated social costs predict symptom improvement in social anxiety disorder? BMC Psychiatry, 19, 80. 10.1186/s12888-019-2054-2 [PubMed: 30795749]

Kang N, Ding D, Van Riemsdijk MB, Morina N, Neerincx MA, & Brinkman W-P (2021). Self-identification with a virtual experience and its moderating effect on self-efficacy and presence. International Journal of Human-Computer Interaction, 37, 181–196. 10.1080/10447318.2020.1812909

Kennedy RS, Lane NE, Berbaum KS, & Lilienthal MG (1993). Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness. The International Journal of Aviation Psychology, 3, 203–220. 10.1207/s15327108ijap0303_3

Kessler RC, Petukhova M, Sampson NA, Zaslavsky AM, Wittchen H, & Ullrich. (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. International Journal of Methods in Psychiatric Research, 21, 169–184. 10.1002/mpr.1359 [PubMed: 22865617]

Kim HE, Hong YJ, Kim MK, Jung YH, Kyeong S, & Kim JJ (2017). Effectiveness of self-training using the mobile-based virtual reality program in patients with social anxiety disorder. Computers in Human Behavior, 73, 614–619. 10.1016/j.chb.2017.04.017

Klinger E, Bouchard S, Légeron P, Roy S, Lauer F, Chemin I, et al. (2005). Virtual reality therapy versus cognitive behavior therapy for social phobia: A preliminary controlled study. CyberPsychology and Behavior, 8, 76–88. 10.1089/cpb.2005.8.76 [PubMed: 15738695]

Krijn M, Emmelkamp PMG, Olafsson RP, & Biemond R (2004). Virtual reality exposure therapy of anxiety disorders: A review. Clinical Psychology Review, 24, 259–281. 10.1016/j.cpr.2004.04.001 [PubMed: 15245832]

Kroenke K, & Spitzer RL (2002). The PHQ-9: A new depression diagnostic and severity measure. Psychiatric Annals, 32, 509–515. 10.3928/0048-5713-20020901-06

Kroenke K, Spitzer RL, & Williams JBW (2001). The PHQ-9. Journal of General Internal Medicine, 16, 606–613. 10.1046/j.1525-1497.2001.016009606.x [PubMed: 11556941]

Kyoto K (2017). Pico Goblin specifications SizeScreens.com, 2017. URL: https://www.sizescreens.com/pico-goblin-specifications/ accessed 2021–03-31.

Leigh E, & Clark DM (2018). Understanding social anxiety disorder in adolescents and improving treatment outcomes: Applying the cognitive model of Clark and Wells (1995). Clinical Child and Family Psychology Review, 21, 388–414. 10.1007/s10567-018-0258-5 [PubMed: 29654442]

Lindner P, Miloff A, Fagernas S, Andersen J, Sigeman M, Andersson G, … Carlbring P (2019). 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. 10.1016/j.janxdis.2018.07.003 [PubMed: 30054173]

Lindner P, Miloff A, Hamilton W, Reuterskiöld L, Andersson G, Powers MB, et al. (2017). Creating state of the art, next-generation virtual reality exposure therapies for anxiety disorders using consumer hardware platforms: Design considerations and future directions. Cognitive Behaviour Therapy, 46, 404–420. 10.1080/16506073.2017.1280843 [PubMed: 28270059]

Lindner P, Miloff A, Zetterlund E, Reuterskiöld L, Andersson G, & Carlbring P (2019). Attitudes toward and familiarity with virtual reality therapy among practicing cognitive behavior therapists: A cross-sectional survey study in the era of consumer VR platforms. Frontiers in Psychology, 10, 176. 10.3389/fpsyg.2019.00176 [PubMed: 30800086]

Ling Y, Nefs HT, Morina N, Heynderickx I, & Brinkman WP (2014). A meta-analysis on the relationship between self-reported presence and anxiety in virtual reality exposure therapy for anxiety disorders. PLoS One, 9, Article e96144. 10.1371/journal.pone.0096144

Lister HA, Piercey CD, & Joordens C (2010). The effectiveness of 3-D video virtual reality for the treatment of fear of public speaking. Journal of CyberTherapy and Rehabilitation, 3, 375–381.

Magnusson K (2018). powerlmm: Power analysis for longitudinal multilevel models. R package version 0.4.0 https://CRAN.R-project.org/package=powerlmm.

Maples-Keller JL, Bunnell BE, Kim S-J, & Rothbaum BO (2017). The use of virtual reality technology in the treatment of anxiety and other psychiatric disorders. Harvard Review of Psychiatry, 25, 103–113. 10.1097/HRP.0000000000000138 [PubMed: 28475502]

Mattick RP, & Clarke JC (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. Behaviour Research and Therapy, 36, 455–470. 10.1016/S0005-7967(97)10031-6 [PubMed: 9670605]

Mayo-Wilson E, Dias S, Mavranezouli I, Kew K, Clark DM, Ades AE, et al. (2014). Psychological and pharmacological interventions for social anxiety disorder in adults: A systematic review and network meta-analysis. The Lancet Psychiatry, 1, 368–376. 10.1016/S2215-0366(14)70329-3 [PubMed: 26361000]

McCarthy J, & Goffin R (2004). Measuring job interview anxiety: Beyond weak knees and sweaty palms. Personnel Psychology, 57, 607–637. 10.1111/j.1744-6570.2004.00002.x

Meyer TJ, Miller ML, Metzger RL, & Borkovec TD (1990). Development and validation of the Penn state worry questionnaire. Behaviour Research and Therapy, 28, 487–495. 10.1016/0005-7967(90)90135-6 [PubMed: 2076086]

Miloff A, Lindner P, Dafgård P, Deak S, Garke M, Hamilton W, … Carlbring P (2019). Automated virtual reality exposure therapy for spider phobia vs. in-vivo one-session treatment: A randomized non-inferiority trial. Behaviour Research and Therapy, 118, 130–140. 10.1016/j.brat.2019.04.004 [PubMed: 31075675]

Morina N, Kampmann I, Emmelkamp P, Barbui C, & Hoppen TH (2021). Meta-analysis of virtual reality exposure therapy for social anxiety disorder. Psychological Medicine, 1–3. 10.1017/S0033291721001690

Mostajeran F, Balci M, Steinicke F, Kuhn S, & Gallinat J (2020). March). The effects of virtual audience size on social anxiety during public speaking. In Proceedings of the 2020 IEEE conference on virtual reality and 3D user Interfaces (pp. 303–312). Atlanta, GA: VR). 10.1109/VR46266.2020.1581528155120.

Newman MG, Kachin KE, Zuellig AR, Constantino MJ, & Cashman-McGrath L (2003). The Social Phobia Diagnostic Questionnaire: Preliminary validation of a new self-report diagnostic measure of social phobia. Psychological Medicine, 33, 623–635. 10.1017/S0033291703007669 [PubMed: 12785464]

Newman MG, Szkodny LE, Llera SJ, & Przeworski A (2011). A review of technology-assisted self-help and minimal contact therapies for anxiety and depression: Is human contact necessary for therapeutic efficacy? Clinical Psychology Review, 31, 89–103. 10.1016/j.cpr.2010.09.008 [PubMed: 21130939]

Nolet K, Corno G, & Bouchard S (2020). The adoption of new treatment modalities by health professionals and the relative weight of empirical evidence in favor of virtual reality exposure versus mindfulness in the treatment of anxiety disorders. Frontiers in Human Neuroscience, 14, 86. 10.3389/fnhum.2020.00086 [PubMed: 32269517]

Norton AR, & Abbott MJ (2016). Self-focused cognition in social anxiety: A review of the theoretical and empirical literature. Behaviour Change, 33, 44–64. 10.1017/bec.2016.2

Otte C (2011). Cognitive behavioral therapy in anxiety disorders: Current state of the evidence. Dialogues in Clinical Neuroscience, 13, 413–421. [PubMed: 22275847]

Padhye NS, Cron SG, Gusick GM, Hamlin SK, & Hanneman SK (2009). Randomization for clinical research: An easy-to-use spreadsheet method. Research in Nursing & Health, 32, 561–566. 10.1002/nur.20341

Page S, & Coxon M (2016). Virtual reality exposure therapy for anxiety disorders: Small samples and no controls? Frontiers in Psychology, 7, 326. 10.3389/fpsyg.2016.00326, 326. [PubMed: 27014137]

Paul M, Bullock K, & Bailenson J (2020). Virtual reality behavioral activation as an intervention for major depressive disorder: Case report. JMIR Ment Health, 7, Article e24331. 10.2196/24331

Pinheiro J, Bates D, DebRoy S, & Sarkar D (2020). nlme: Linear and Nonlinear mixed effects models. R package version 3.1 (p. 150). R Core Team https://CRAN.R-project.org/package=nlme.

R Core Team. (2020). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Regenbrecht H, & Schubert T (2002). Real and illusory interactions enhance presence in virtual environments. Presence: Teleoperators and Virtual Environments, 11, 425–434. 10.1162/105474602760204318

Robins LN, Wing J, Wittchen HU, Helzer JE, Babor TF, Burke J, … Towle LH (1988). The Composite International Diagnostic Interview: An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. Archives of General Psychiatry, 45, 1069–1077. 10.1001/archpsyc.1988.01800360017003 [PubMed: 2848472]

Rubin DB (1987). Multiple imputation for nonresponse in surveys. New York: Wiley.

Rupp C, Doebler P, Ehring T, & Vossbeck-Elsebusch AN (2017). Emotional processing theory put to test: A meta-analysis on the association between process and outcome measures in exposure therapy. Clinical Psychology & Psychotherapy, 24, 697–711. 10.1002/cpp.2039 [PubMed: 27561691]

Safir MP, Wallach HS, & Bar-Zvi M (2012). Virtual reality cognitive-behavior therapy for public speaking anxiety: One-year follow-up. Behavior Modification, 36, 235–246. 10.1177/0145445511429999 [PubMed: 22180390]

Schmidt NB, Keough ME, Hunter LR, & Funk AP (2008). Physical illness and treatment of anxiety disorders: A review. In Zvolensky MJ, & Smits JAJ (Eds.), Anxiety in health behaviors and physical Illness (pp. 341–366). New York, NY: Springer Science. 10.1007/978-0-387-74753-8_13.

Schubert T, Friedmann F, & Regenbrecht H (2001). The experience of presence: Factor analytic insights. Presence: Teleoperators and Virtual Environments, 10, 266–281. 10.1162/105474601300343603

Sheehan DV, Lecrubier Y, Harnett Sheehan K, Janavs J, Weiller E, Keskiner A, et al. (1997). The validity of the Mini international neuropsychiatric interview (MINI) according to the SCID-P and its reliability. European Psychiatry, 12, 232–241. 10.1016/S0924-9338(97)83297-X

Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. Journal of Clinical Psychiatry, 59(Suppl. 20), 22–33.

Stangier U (2016). New developments in cognitive-behavioral therapy for social anxiety disorder. Current Psychiatry Reports, 18, 25. 10.1007/s11920-016-0667-4 [PubMed: 26830883]

Starcevic V, Berle D, Milicevic D, Hannan A, Lamplugh C, & Eslick GD (2007). Pathological worry, anxiety disorders and the impact of co-occurrence with depressive and other anxiety disorders. Journal of Anxiety Disorders, 21, 1016–1027. 10.1016/j.janxdis.2006.10.015 [PubMed: 17270391]

Stolz T, Schulz A, Krieger T, Vincent A, Urech A, Moser C, … Berger T (2018). A mobile app for social anxiety disorder: A three-arm randomized controlled trial comparing mobile and PC-based guided self-help interventions. Journal of Consulting and Clinical Psychology, 86, 493–504. 10.1037/ccp0000301 [PubMed: 29781648]

Summers BJ, Schwartzberg AC, & Wilhelm S (2021). A virtual reality study of cognitive biases in body dysmorphic disorder. Journal of Abnormal Psychology, 130, 26–33. 10.1037/abn0000563 [PubMed: 33252917]

Teo AR, Lerrigo R, & Rogers MAM (2013). The role of social isolation in social anxiety disorder: A systematic review and meta-analysis. Journal of Anxiety Disorders, 27, 353–364. 10.1016/j.janxdis.2013.03.010 [PubMed: 23746493]

Vasconcelos-Raposo J, Bessa M, Melo M, Barbosa L, Rodrigues R, Teixeira CM, … Sousa AA (2016). Adaptation and validation of the igroup presence questionnaire (IPQ) in a Portuguese sample. Presence, 25, 191–203. 10.1162/PRES_a_00261

Vrielynck N, & Philippot P (2009). Regulating emotion during imaginal exposure to social anxiety: Impact of the specificity of information processing. Journal of Behavior Therapy and Experimental Psychiatry, 40, 274–282. 10.1016/j.jbtep.2008.12.006 [PubMed: 19168173]

Vriends N, Bolt OC, & Kunz SM (2014). Social anxiety disorder, a lifelong disorder? A review of the spontaneous remission and its predictors. Acta Psychiatrica Scandinavica, 130, 109–122. 10.1111/acps.12249 [PubMed: 24506164]

Wallach HS, Safir MP, & Bar-Zvi M (2011). Virtual reality exposure versus cognitive restructuring for treatment of public speaking anxiety: A pilot study. Israel Journal of Psychiatry and Related Sciences, 48, 91–97.

Whiteside SPH, Brennan E, Biggs BK, Vickers K, Hathaway J, Seifert SJ, … Hofschulte DR (2020). The feasibility of verbal and virtual reality exposure for youth with academic performance worry. Journal of Anxiety Disorders, 76, 102298. 10.1016/j.janxdis.2020.102298 [PubMed: 32937260]

Wolpe J, & Lazarus AA (1967). Behavior therapy techniques: A guide to the treatment of neuroses. Oxford: Pergamon Press.

Wong QJJ, Gregory B, McLellan LF, Kangas M, Abbott MJ, Carpenter L, et al. (2017). Anticipatory processing, maladaptive attentional focus, and postevent processing for interactional and performance situations: Treatment response and relationships with symptom change for individuals with social anxiety disorder. Behavior Therapy, 48, 651–663. 10.1016/j.beth.2017.03.004 [PubMed: 28711115]

Zalta AK, & Foa EB (2012). Exposure therapy: Promoting emotional processing of pathological anxiety Cognitive Behavior Therapy. 10.1002/9781118470886.ch4

```
                    Assessed for eligibility (n = 49)
                                │
                                ├──────────────────►  Excluded (n = 5)
                                │                     Did not meet DSM-5 diagnostic criteria for SAD (n = 5)
                                ▼
                        Randomized (n = 44)
                                │
                ┌───────────────┴───────────────┐
                ▼                               ▼
    VR Exposure Therapy (n = 26)            Waitlist (n = 18)
    Completed initial assessment (n = 26)   Completed initial assessment (n = 18)
    Completed 4-10 sessions during the      Completed post-treatment assessment (n = 16)
      first phase (n = 23)                  Dropout before treatment (n = 2)
    Completed post-treatment assessment
      (n = 23)                                          │
    Dropout during treatment (n = 3; all               ▼
      due to COVID-19 )              VR Exposure Therapy After Waitlist (n = 16)
                │                    Completed 4-10 sessions during the first phase (n = 15)
                │                    Dropout during treatment (n = 1)
                │                    Completed 3-month follow-up assessment (n = 15)
                ▼                    Completed 6-month follow-up assessment (n = 14)
    Completed 3-month follow-up      Dropout at follow-up (n = 1)
      assessment (n = 23)
    Completed 6-month follow-up                        │
      assessment (n = 21)                              │
    Dropout at follow-up (n = 2)                       │
                │                                      ▼
                ▼
    Intention-to-treat sample (n = 26)   Intention-to-treat sample (n = 18)
    Completer sample (n = 21)            Completer sample (n = 14)
```

Note. CONSORT = Consolidated Standards of Reporting Trials; VRE = virtual reality exposure therapy; SAD = social anxiety disorder.

Fig. 1.
CONSORT flow chart

Note. VRE = virtual reality exposure therapy. Participant introduced themselves to the host with a party guide present.

Fig. 2a.
Scene 1 in the Dinner Party-Themed VRE Protocol for General Social Anxiety

Note. VRE = virtual reality exposure therapy. Participant engaged in a group format interview with two other interviewees.

Fig. 2b.
Scene 6 in the Job Interview Anxiety-Themed VRE Protocol

Table 1

Descriptive statistics of outcome measures at all time-points.

		Pre-Treatment		Post-Treatment		3MFU		6MFU	
		M	(SD)	M	(SD)	M	(SD)	M	(SD)
SAD composite	WL	0.88	(0.84)	−0.09	(0.89)	–	–	–	–
	VRE	1.31	(0.61)	−0.52	(0.60)	−0.45	(0.46)	−0.49	(0.77)
SPDQ	WL	16.99	(4.42)	12.53	(4.92)	–	–	–	–
	VRE	19.12	(3.50)	9.52	(3.51)	10.10	(2.63)	10.26	(4.30)
SIAS	WL	46.45	(14.54)	29.80	(13.28)	–	–	–	–
	VRE	53.48	(9.53)	25.29	(9.50)	25.82	(8.68)	23.95	(12.00)
MASI	WL	70.46	(16.27)	47.89	(22.48)	–	–	–	–
	VRE	74.87	(15.80)	38.35	(14.72)	42.21	(12.16)	37.80	(18.08)
PSWQ	WL	60.19	(7.81)	47.09	(7.71)	–	–	–	–
	VRE	67.60	(6.92)	46.42	(8.73)	49.01	(10.06)	43.01	(15.12)
PHQ-9	WL	11.20	(6.01)	7.94	(3.60)	–	–	–	–
	VRE	13.20	(5.80)	6.46	(3.11)	7.46	(1.57)	6.22	(4.07)

Note. FU = follow-up; MASI = job interview anxiety scale; PHQ-9 = patient health questionnaire–9-item; PSWQ = Penn State worry questionnaire; SAD composite = averaged standardized score of the social interaction anxiety scale and social phobia diagnostic questionnaire; VRE = virtual reality exposure therapy; WL = waitlist control condition.

Table 2

Within-group simple slopes analysis of multilevel modeling with random intercept findings.

		Pre-post		Pre-3MFU		Pre-6MFU		Post-3MFU		Post-6MFU	
		β	g	β	g	β	g	β	G	β	g
SAD	WL	**-0.86****	**-2.35**	–	–	–	–	–	–	–	–
Composite	VRE	**-1.84*****	**-4.77**	**-0.60*****	**-4.80**	**-0.87*****	**-3.86**	0.01	0.12	-0.04	-0.13
SPDQ	WL	**-3.83****	**-1.93**	–	–	–	–	–	–	–	–
	VRE	**-9.59*****	**-4.31**	**-4.50*****	**-4.50**	**-2.92*****	**-3.30**	0.37	0.47	0.16	0.07
SIAS	WL	**-15.34*****	**-2.50**	–	–	–	–	–	–	–	–
	VR	**-28.12*****	**-4.58**	**-13.75*****	**-4.60**	**-9.78*****	**-4.04**	-0.67	-0.38	-1.87	-0.34
MASI	WL	**-21.59*****	**-2.51**	–	–	–	–	–	–	–	–
	VRE	**-36.28*****	**-4.17**	**-16.33*****	**-3.70**	**-12.35*****	**-3.77**	-0.28	-0.10	-4.41	-0.53
PSWQ	WL	**-13.17*****	**-2.80**	–	–	–	–	–	–	–	–
	VRE	**-21.28*****	**-4.45**	**-9.24*****	**-3.27**	**-8.19*****	**-3.28**	-1.71	-0.65	-6.00	-0.66
PHQ-9	WL	-2.79	**-1.09**	–	–	–	–	–	–	–	–
	VRE	**-6.70*****	**-2.21**	**-2.88*****	**-2.21**	**-2.28*****	**-2.00**	-0.12	-0.15	-1.24	-0.51

β = regression weight; g = Hedge's g effect size; FU = follow-up; MASI = job interview anxiety scale; PHQ-9 = patient health questionnaire–9-item; Pre-3MFU = pre-treatment to 3-month follow-up; Pre-6MFU = pre-treatment to 6-month follow-up; Post-3MFU = pre-treatment to 3-month follow-up; PSWQ = Penn State worry questionnaire; SAD composite = averaged standardized score of the social interaction anxiety scale and social phobia diagnostic questionnaire; VRE = virtual reality exposure therapy.

*
p < .05

p < .01

p < .001.

Table 3

Multilevel model with random intercepts for pre-to post-treatment time-point, condition, and their interaction predicting outcome measures.

		Parameter Estimates			
		β	**(SE)**	*t*	*g*
SAD Composite	Intercept	**0.90** ***	**(0.17)**	**5.39**	**1.61**
	Pre-Post	**−0.91** ***	**(0.21)**	**−4.29**	**−1.32**
	Condition	0.40	(0.21)	1.90	0.58
	Pre-Post × Condition	**−0.97** **	**(0.28)**	**−3.45**	**−1.06**
SPDQ	Intercept	**17.02** ***	**(0.93)**	**18.31**	**5.49**
	Pre-Post	**−4.11** **	**(1.18)**	**−3.48**	**−1.07**
	Condition	2.08	(1.19)	1.75	0.54
	Pre-Post × Condition	**−5.73** ***	**(1.58)**	**−3.63**	**−1.11**
SIAS	Intercept	**46.80** ***	**(2.67)**	**17.50**	**5.24**
	Pre-Post	**−15.90** ***	**(3.45)**	**−4.61**	**−1.41**
	Condition	6.45	(3.44)	1.87	0.57
	Pre-Post × Condition	**−12.81** **	**(4.59)**	−2.79	**−0.85**
MASI	Intercept	**70.70** ***	**(3.92)**	**18.05**	**5.41**
	Pre-Post	**−22.57** ***	**(4.83)**	**−4.68**	**−1.43**
	Condition	4.01	(4.98)	0.81	0.25
	Pre-Post × Condition	**−13.95** *	**(6.48)**	−2.15	**−0.66**
PSWQ	Intercept	**60.21** ***	**(1.84)**	**32.66**	**9.79**
	Pre-Post	**−12.89** ***	**(2.54)**	**−5.08**	**−1.56**
	Condition	**7.37****	**(2.40)**	**3.08**	**0.94**
	Pre-Post × Condition	**−8.43** *	**(3.33)**	−2.53	**−0.78**
PHQ-9	Intercept	**11.19** ***	**(1.13)**	**9.94**	**2.98**
	Pre-Post	**−3.19** *	**(1.57)**	**−2.03**	**−0.62**
	Condition	2.03	(1.46)	1.39	0.42
	Pre-Post × Condition	−3.61	(2.05)	−1.76	−0.54

β = regression weight; *g* = Hedge's *g* effect size; MASI = job interview anxiety scale; PHQ-9 = patient health questionnaire–9-item; PSWQ = Penn State worry questionnaire; SAD composite = averaged standardized score of the social interaction anxiety scale and social phobia diagnostic questionnaire; SE = standard error; *t* = *t*-statistic for the specific parameter estimate.

*
p < .05

**
p < .01

***
p < .001.

Table 4

Comparison presence and cybersickness with prior data.

Domains of Presence	Current study (n = 44)		Regenbrecht and Schubert (2002)		Mostajeran et al. (2020)		Current study vs. Regenbrecht and Schubert (2002)		Current study vs. Mostajeran et al. (2020)	
	M	(*SD*)	(n = 56) *M*	(*SD*)	(n = 24) *M*	(*SD*)	*p*	*g*	*p*	*g*
Spatial presence	3.49	(0.81)	2.92	(1.16)	4.02	(3.80)	.002	0.56	.513	−0.19
Involvement	3.84	(1.05)	3.38	(1.42)	3.43	(2.26)	.062	0.37	.892	0.04
Realness	2.67	(0.75)	1.86	(1.15)	3.13	(1.90)	.008	0.83	.373	0.25

Cybersickness	Current study (n = 44)		Bouchard et al. (2009) (n = 157)		Kim et al. (2017) (n = 52)		Current study vs. Bouchard et al. (2009)		Current study vs. Kim et al. (2017)	
	M	(*SD*)	*M*	(*SD*)	*M*	(*SD*)	*p*	*g*	*p*	*g*
Total score	16.39	(11.51)	36.27	(31.46)	16.10	(15.67)	<.001	−0.84	.918	0.02

Note. Presence was measured using the iGroup Presence Questionnaire (IPQ), and cybersickness was assessed with the Simulation Sickness Questionnaire (SSQ).