Behaviour Research and Therapy 118 (2019) 130–140

Contents lists available at ScienceDirect

Behaviour Research and Therapy

journal homepage: www.elsevier.com/locate/brat

Automated virtual reality exposure therapy for spider phobia vs. in-vivo one-session treatment: A randomized non-inferiority trial

Alexander Miloff[a,*], Philip Lindner[a,b], Peter Dafgård[a], Stefan Deak[a], Maria Garke[b], William Hamilton[c], Julia Heinsoo[a], Glenn Kristoffersson[a], Jonas Rafi[a], Kerstin Sindemark[a], Jessica Sjölund[a], Maria Zenger[a], Lena Reuterskiöld[a], Gerhard Andersson[b,d], Per Carlbring[a,e]

[a] Department of Psychology, Stockholm University, Stockholm, Sweden
[b] Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet & Stockholm Health Care Services, Stockholm County, Sweden
[c] Mimerse AB, Stockholm, Sweden
[d] Department of Behavioral Sciences and Learning, Linköping University, Linköping, Sweden
[e] Department of Psychology, University of Southern Denmark, Denmark

ARTICLE INFO

Keywords:
Exposure therapy
One-session treatment
Virtual reality
Spider phobia

ABSTRACT

Objective: This study compared the efficacy of a technician-assisted single-session virtual reality exposure therapy (VRET) for the treatment of spider phobia featuring low-cost consumer-available hardware and novel automated software to gold-standard in-vivo one-session treatment (OST), using a parallel group randomized non-inferiority design. *Method* Participants (N = 100) were randomized to VRET and OST arms. Assessors blinded to treatment allocation evaluated participants at pre- and post-treatment as well follow-up (3 and 12 months) using a behavioral approach test (BAT) and self-rated fear of spider, anxiety, depression and quality-of-life scales. A maximum post-treatment difference of 2-points on the BAT qualified as non-inferiority margin. *Results* Linear mixed models noted large, significant reductions in behavioral avoidance and self-reported fear in both groups at post-treatment, with VRET approaching the strong treatment benefits of OST over time. Non-inferiority was identified at 3- and 12- months follow-up but was significantly worse until 12-months. There was no significant difference on a questionnaire measuring negative effects. *Conclusions* Automated VRET efficaciously reduced spider phobia symptoms in the short-term and was non-inferior to in-vivo exposure therapy in the long-term. VRET effectiveness trials are warranted to evaluate real-world benefits and non-specific therapeutic factors accruing from the presence of a technician during treatment. ClinicalTrials.gov (NCT02533310).

1. Introduction

Specific phobia is a common disorder (Wardenaar et al., 2017), characterized by excessive and persistent fear of object or situation (American Psychological Association, 2013). Symptoms including anxiety, worry and avoidance have negative repercussions for work, leisure and quality of life (Choy, Fyer, & Lipsitz, 2007), and may predict the onset of more serious conditions (Lieb et al., 2016). Despite decades long availability of highly effective treatments for specific phobia (Barlow, 2002), particularly in-vivo exposure emphasizing direct, systematic contact with feared stimuli and showing strong effect sizes relative to wait-list ($d = 1.05$) and placebo ($d = 0.68$; Wolitzky-Taylor, Horowitz, Powers, & Telch, 2008), dissemination of this evidence-based treatment has been lacking (Deacon et al., 2013).

Virtual reality exposure therapy (VRET) is a technological approach to providing patients the experience of interacting therapeutically with feared stimuli and situations. Using head-mounted displays, gyroscopes and accelerometers to track movement, computer generated three-dimensional environments are projected to users along with sound and dynamically adjusted in response to their natural head, hand or other motion. Given sufficient synchrony between user movement, visual display, and dynamic changes caused by or to the user, the illusion of place (being present in a virtual space) and plausibility (that events are really happening) can be created ('presence'; Freeman et al., 2017). Advantages to VRET over in-vivo exposure therapy include ensuring therapy can be conducted in an office, safeguarding ethical standards such as confidentiality and patient-client boundaries (Olatunji, Deacon, & Abramowitz, 2009), minimizing barriers to treatment with improved acceptability (Garcia-Palacios, Botella, Hoffman, & Fabregat, 2007), and in the future, reduced costs and improved geographic access using

* Corresponding author. Department of Psychology, Stockholm University, 106 91, Stockholm, Sweden.
E-mail address: alexander.miloff@psychology.su.se (A. Miloff).

https://doi.org/10.1016/j.brat.2019.04.004
Received 15 June 2018; Received in revised form 23 January 2019; Accepted 9 April 2019
Available online 16 April 2019
0005-7967/ © 2019 Published by Elsevier Ltd.

automated self-help VRET treatments not requiring a therapist (Newman, Szkodny, Llera, & Przeworski, 2011).

Early VRET treatments for fear of spiders, among the most common specific phobias (Oosterink, De Jongh, & Hoogstraten, 2009), have shown large reductions in behavioral avoidance and/or self-reported fear in clinical populations (Bouchard, Côté, St-Jacques, Robillard, & Renaud, 2006; Carlin, Hoffman, & Weghorst, 1997; Côté & Bouchard, 2005; Garcia-Palacios, Hoffman, Carlin, Furness, & Botella, 2002; Hoffman, Garcia-Palacios, Carlin, Furness, & Botella-Arbona, 2003). Using a single-session format and short (30–45 min) exposures, more recent VRET studies have also noted significant effects but these have been smaller (Kleim et al., 2014; Minns et al., 2018). Two randomized controlled trials comparing in-vivo exposure therapy to VRET found no significant difference in behavioral avoidance at post-measure and follow-up in one study (Michaliszyn, Marchand, Bouchard, Martel, & Poirier-Bisson, 2010) and using a child-population (with unequal allocation at randomization), a moderate advantage for in-vivo treatments at post-measure in another (St-Jacques, Bouchard, & Bélanger, 2010). According to a series of meta-analyses, large effect size improvements and/or no significant difference between VRET and in-vivo exposure therapy have been identified for a number of specific phobia diagnoses (Morina, Ijntema, Meyerbröker, & Emmelkamp, 2015; Opriş et al., 2012; Parsons & Rizzo, 2008; Carl et al., 2019). Deterioration rates among patients receiving VR treatment for anxiety disorders are similar to other active treatments according to an individual patient data-level meta-analysis (4 vs 2.5%) and lower than wait-list controls (15%; Fernández-Álvarez et al., 2018).

Evidence in favor of VRET notwithstanding, access to this treatment has been limited by historically slow, expensive and cumbersome hardware (Turner & Casey, 2014), treatments that normally required multiple visits to conduct (Mühlberger, Weik, Pauli, & Wiedemann, 2006), software developed for use in specialist laboratories, and interventions that replaced in-vivo stimuli with virtual but required a human therapist to administer (Freeman et al., 2017). Criticism has also been made of study quality (Page & Coxon, 2016; Turner & Casey, 2014), such as small sample sizes (McCann et al., 2014), inadequate null hypothesis significance testing (Parsons, 2015), lack of long-term follow-ups and high drop-out rates (Valmaggia, Latif, Kempton, & Rus-Calafell, 2016).

The goal of the current study is to compare in a clinical sample of spider phobic patients, the efficacy of an inexpensive consumer-available virtual reality device and automated VRET application with technician-assistance (Lindner et al., 2017), to one-session treatment (OST), a brief, modern and intensive form of in-vivo exposure therapy conducted during a single massed session (Öst, 1989). Treatment components of OST include graded in-vivo exposure, participant modeling, reinforced practice, and systematically identifying and eliminating catastrophic cognitions (Öst, 2012). Past research has suggested that therapist directed OST is efficacious for reducing both behavioral avoidance ($d = 2.73$; Andersson et al., 2009) and self-reported fear ($d = 2.84$ to 3.05; Zlomke & Davis III, 2008) in spider phobic patients, and with results comparable to treatments lasting hours longer.

Given the established evidence-base for VRET, indicating no significant difference with traditional in-vivo exposure therapy in adults, and the ethical concerns of withholding a known effective treatment, this study was conducted as a non-inferiority trial. Non-inferiority trials are the preferred design for evaluating difference between a novel intervention and reference treatment when non-significant difference is assumed (Piaggio, Elbourne, Altman, Pocock, & Evans, 2012). We hypothesized that this novel single-session VRET would be not inferior to in-vivo exposure therapy according to a 2-point behavioral approach test non-inferiority margin (i.e., minimum clinically significant change) at post-assessment and long-term follow-up, with similar low levels of negative treatment effects.

2. Method

This study was designed as a parallel group randomized non-inferiority trial. A research protocol was published (Miloff et al., 2016) and the trial pre-registered in the ClinicalTrials.gov database (NCT02533310). Results are published in accordance with SPIRIT (Chan et al., 2013) and CONSORT guidelines for randomized controlled (Schulz, Altman, & Moher, 2010) and non-inferiority trials (Piaggio et al., 2012). Ethical approval for this study was granted by the Stockholm Regional Ethical Review Board (2015/472-31).

2.1. Participants

In order to recruit a diverse sample (Lindner, Nyström, Hassmén, Andersson, & Carlbring, 2015), advertising was conducted across a variety of online and print media. Recruitment was conducted between August and December 2015, when the last subject was randomized. Eligibility criteria for participation was: a) primary diagnosis of spider phobia (American Psychological Association, 2013) and, b) a score of 9 points or less on an in-vivo spider behavioral approach test (BAT; see below). Participants were also required to: c) be ≥ 18 years, d) resident of Sweden, e) speak fluent Swedish, f) agree to be randomized to one of two treatment groups, and g) travel to Stockholm University on at least three separate occasions for treatment and assessment. A fourth assessment occasion (12-month follow-up) was added after participant inclusion and following an amendment to the ethical approval. Exclusion criteria included: a) suicidal ideation, b) substance abuse, c) serious ongoing/concurrent mental disorders such as bipolar disorder and psychosis, d) ongoing psychotherapy or psychotropic treatment (unless dosage stable ≥ 3 months and no planned changes during study duration), and e) no eye-sight or balance problems (such as lack of stereoscopic vision) that could impair the virtual reality experience. Use of prescription glasses or contact lenses was acceptable.

A total of 217 participants responded to advertisements and 188 completed an online screening battery (see Fig. 1). Of these, 68 were excluded for failing to meet inclusion criteria. The remaining participants were invited to Stockholm University in order of registration and availability to complete an in-person BAT, diagnostic interview and self-report battery using Swedish translations of pre-existing and novel questionnaires. Pre-assessments were completed by 107 participants (n = 7 excluded) in order to fulfill the preset target of 100 participants randomized to OST or VRET. Demographic characteristics of the sample are described in Table 1 with no significant differences between them (all $p > .55$). Mean age was M = 34.05 (SD = 10.35) and primarily female (> 80%). Baseline primary outcome measures were similar to previous OST studies such as Öst, Ferebee, and Furmark (1997).

2.2. Measures

2.2.1. Primary outcome measure

An established Behavioral Approach Test (BAT; Öst, Salkovskis, & Hellström, 1991) featuring a 13-point scale (scored 0–12) measuring in-vivo spider avoidance was used as the primary outcome measure. Outside the doorway of identical dedicated rectangular (3 × 5 m) laboratory rooms, participants were given standardized instructions to enter and approach a desk on the far side of the room on which sat a spider enclosed in a transparent plastic container with lid (40 × 30 × 19 cm). Spiders were medium sized harmless varieties (such as *Tegenaria* and *Araneidae*) native to Sweden and the same approx. 2 cm diameter (including legs) as used in Ollendick et al. (2009). Participants were instructed that the aim of the test was to approach the spider, pick it up and hold it for 20 s. They were encouraged to do their best but were told they could abort the test at any time. Points were given in accordance with standardized approach steps (i.e., higher score equals less avoidance; see Miloff et al., 2016). The BAT was completed at pre-assessment, post-assessment and 3- and 12-month follow-up.

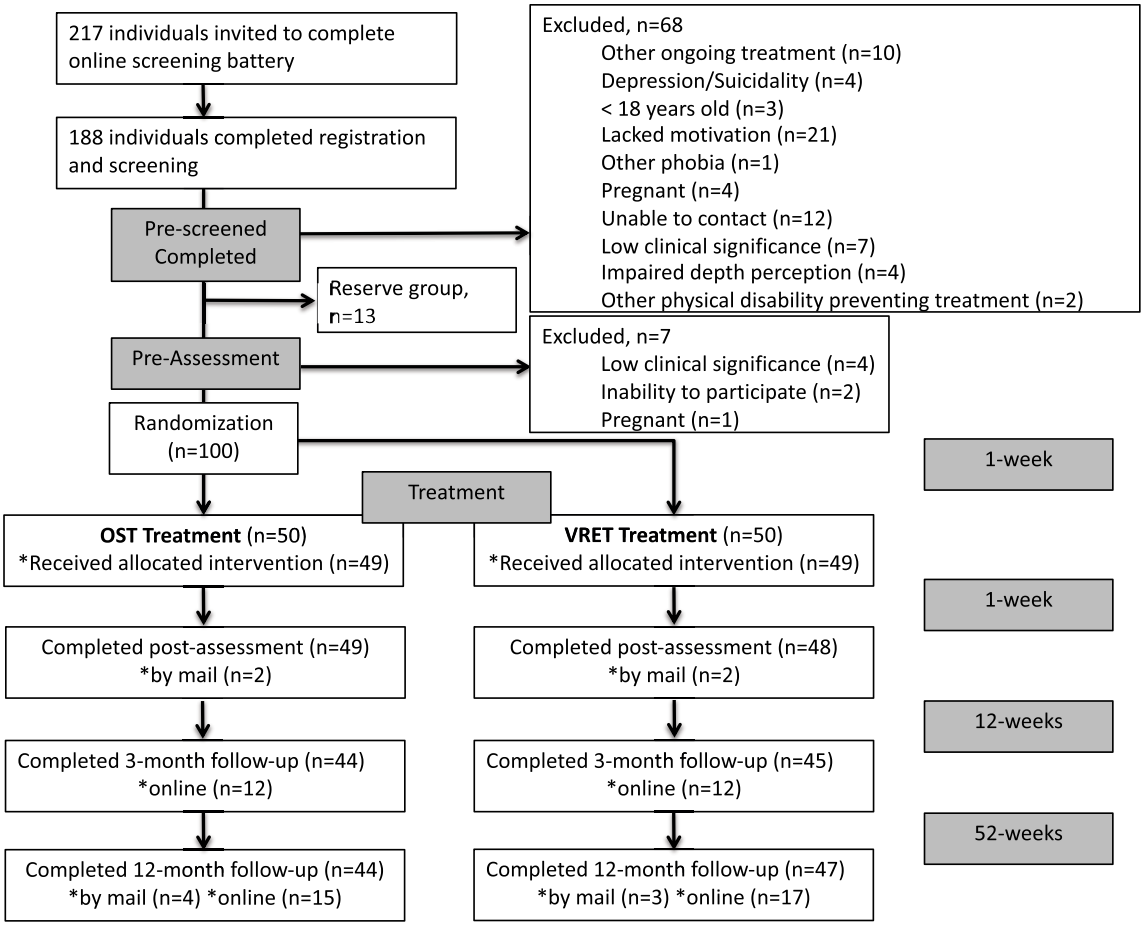

Fig. 1. Flow of participants through study. OST = One-Session Treatment; VRET = Virtual Reality Exposure Therapy.

2.2.2. Secondary outcome measures

2.2.2.1. Specific phobia diagnosis. The Structured Clinical Interview for DSM-IV Axis-I Disorders (SCID-I/P) (First, Robert, Gibbon, & Williams, 2002), adapted for DSM-5 criteria (American Psychological Association, 2013) was used to provide clinical diagnosis for spider phobia. All interviewers were trained and examined in administration of the SCID as part of their coursework and completed study-specific training in SCID administration (including scoring examples). Although reliability data was not collected, SCID-I/P inter-rater kappa statistic has been found to be high in the case of specific phobia (kappa = .80; Lobbestael, Leurgans, & Arntz, 2011). Initial specific phobia diagnostic interviews were conducted prior to randomization and all post-treatment interviewers (3- and 12-month follow-up) were blinded to treatment allocation.

2.2.2.2. Self-rated phobia symptoms. The Spider Phobia Questionnaire (SPQ) is a 31-item scale measuring fear of spiders according to true/false scenarios such as "I dislike looking at pictures of spiders in a magazine" (Klorman, Weerts, Hastings, Melamed, & Lang, 1974). The Fear of Spiders Questionnaire (FSQ) features 18 questions about fear and avoidance of spiders (Szymanski & O'Donohue, 1995). Questions such as "If I saw a spider now, I would think it will harm me" are rated on a 7-point scale (1 = *not at all*, 7 = *very much*). The FSQ uniquely captures the construct "fear of harm" and has improved time scale discrimination as compared to the SPQ (Muris & Merckelbach, 1996). Higher scores on both questionnaires relate to greater fear of spiders. The SPQ and FSQ were administered at all assessment occasions including pre-screening.

2.2.3. Other measures

Three self-report measures were used to determine depression, anxiety and quality of life during the clinical trial. The Generalized Anxiety Disorder Assessment (GAD-7; Spitzer, Kroenke, Williams, & Lo, 2006) 7-item questionnaire was used as a generic measure of anxiety symptoms. The 9-item Patient Health Questionnaire (PHQ-9; Kroenke & Spitzer, 2002) was used to measure depressive symptoms. Higher scores on the PHQ and GAD reflect worse health. The Brunnsviken Brief Quality of Life Inventory (BBQ; Lindner, Frykheden, Forsström, Andersson, Ljótsson, 2016) is a measure of subjective quality of life across six empirically derived domains of life. Higher scores on the BBQ reflects better health. These self-rating scales were completed at all assessment occasions including pre-screening.

The Negative Effects Questionnaire (NEQ-32; Rozental, Kottorp, Boettcher, Andersson, & Carlbring, 2016) was used to capture patient self-reported experience of unwanted and adverse events related to their psychotherapy. Items such as "I experienced more hopelessness" were reported based on a 5-level scale of "how negatively it affected me" from (0) *Not at all* to (4) *Extremely*. The NEQ was completed at post-measurement.

The Igroup Presence Questionnaire (IPQ; Schubert, Friedmann, & Regenbrecht, 2001) measures self-reported experience of being present in a virtual environment ("presence"). The questionnaire is made up of 14 items such as "in the computer generated world I had a sense of 'being there,'" measured on a 7-point scale (−3 = *do not agree at all*, 3 = *agree entirely*). The questionnaire is composed of three presence subscales, including spatial presence (sense of being there), involvement (attention within the virtual world), and realness (comparability to reality). All reverse-coded items were compiled so that positive numbers indicated increased presence. Item 6 was removed from

Table 1
Demographic characteristics of participants randomized to treatment groups.

	One-Session Treatment		Virtual Reality Exposure Therapy		
Variables	n (mean)	% (SD)	n (mean)	% (SD)	df, X^2
Age	50 (34.04)	100 (9.85)	50 (34.06)	100 (10.92)	
Gender					2, 0.603
Male	8	16	8	16	
Female	41	82	42	84	
Other	1	2	0	0	
Education					3, 0.168
Grade-school	3	6	0	0	
High school/Trade school	18	36	15	30	
University or college	28	56	35	70	
Doctoral studies	1	2	0	0	
Marital status					2, 0.584
Single	11	22	15	30	
Married, registered partner	37	74	34	68	
Other	2	4	1	2	
Children					1, 0.838
Yes	20	40	19	38	
Previous psychological treatment					1, 0.476
Yes	13	26	10	20	
Previous psychotropic treatment					1, 0.505
Yes	6	12	4	8	
Currently on sick leave					1, 0.079
Yes	3	6	0	0	

analysis due to a typo in printing. The IPQ was completed by VRET participants at post-measure.

2.3. Treatment

2.3.1. One-session treatment (OST)

In-vivo exposure therapy was completed one-on-one with participants during a single up to 3-hr (M = 171.52 min; SD = 30.71) session according to an established treatment manual (Öst, 2012). Psychologists and student psychologists with clinical CBT experience in at least the last year of their 5-year training conducted all OST sessions. Training was provided by an experienced licensed psychologist and psychotherapist, and supervision conducted by the same individual on at least one occasion in groups.

After initial psychoeducation, treatment consisted of gradated and repeated systematic exposures to small (5–15 mm), medium (15–25 mm) and large (> 25 mm) spiders. Prior to moving to a new phase of treatment (e.g., larger spider), participants were asked to report their subjective unit of distress (SUD) on a scale of 0–100 (100 reflecting the most anxiety-provoking experience they had in relation to the phobia). Participants were encouraged to treat interactions with spiders as behavioral experiments, reflecting on catastrophic beliefs, exploring what might occur at each stage of treatment and noting violations of expectancy. Breaks such as for lavatory access were permitted but were otherwise not encouraged. Safety behaviors were discussed and monitored with participants and healthy non-phobic behavior modeled. Treatment ended when catastrophic cognitions and SUD ratings were down to zero (0–100) or as close as possible in the given session of 3 h.

2.3.2. Virtual reality exposure therapy (VRET)

2.3.2.1. Hardware. The hardware used in this trial was the Samsung Gear VR system, comprising one of two Samsung smartphones (Note 4 or Galaxy S6), and the Samsung Gear VR headset (Second Innovator Edition). Unlike traditional virtual reality equipment which are tethered to a computer and require external base stations to track horizontal and vertical movement, the Samsung Gear VR is a mobile system, can run on battery power for multiple hours, and tracks head-rotation using internal sensors and gyroscopes. The Samsung Gear VR headset includes a dedicated touchpad for interacting with virtual reality content, back button for exiting or pausing levels, volume buttons, and a focus adjustment wheel to allow users to sharpen the image depending on their eye-sight. Inexpensive over the ear headphones were provided to deliver verbal instructions and sound.

2.3.2.2. Software design and treatment procedure. The VRET application, called VIMSE, was developed with many of the therapeutic characteristics of OST in mind (Lindner et al., 2017). Intended to be completed in roughly 3-hr time (M = 142.82 min; SD = 21.85), the application once started was fully automated. Therapists, however were present as a "computer technician" to help start the application, address technical problems, ensure treatment compliance, and (for ethical reasons) provide care in case of a serious emotional response. Patients were instructed that they could remove the headset at any time during the treatment if they needed a break or the treatment was overwhelming. They were also informed that the Escape button on the headset could be used to pause treatment. The participant had full-control over how quickly they progressed through the 8 levels of increasingly realistic spiders. Similar to in-vivo exposure sessions, SUD ratings (0–100) were requested by the application using a popup menu at each stage of therapy. Throughout treatment a virtual therapist and spider expert accompanied the participant, delivering voiceover psychoeducation about phobias, supportive feedback regarding treatment progress and information on the importance of spiders, their biology and life cycle. Unlike traditional OST, the VRET application was created as a serious game (Botella et al., 2011). Game playing elements described within the study research protocol (Miloff et al., 2016), included points to collect, puzzles to solve, levels to overcome and other simple game mechanics.

After the data collection period, all VRET technicians completed a short survey about the type of support provided during the automated VRET treatment. Of the 49 VRET treatments, 13 (27%) participants were provided basic technician-level support, such as affirmations ("just keep going, it will be OK!) or suggestions ("do you want to take a break?"). In the case of 5 (10%) of these patients, additional therapist-level support requiring a knowledge of specific phobia and exposure therapy was provided (e.g., discussing catastrophic thoughts).

2.4. Procedure

Participants were invited to participate in the trial using a secure web-based platform (Iterapi) to collect data at pre-screening and other assessment periods if the participant was unable to attend in person (Vlaescu, Alasjö, Miloff, Carlbring, & Andersson, 2016). Randomization was completed by a researcher not otherwise involved in the study using an allocation table created by a computerized random number generator (www.random.org) in blocks of 4, 6, 8 or 10 and with even group allocation in each block. Allocation was concealed until requested by the second author, who enrolled all participants.

At the pre-assessment occasion, participants were provided written informed consent. Approximately 1-week following pre-assessment, participants returned to Stockholm University where they learned of their allocation and completed the roughly 3-hr long treatment. The treatment experience of both groups was intended to be matched on all aspects except the intervention itself. Irrespective of randomization, participants were met in the lobby of the university building, provided basic information about what they would be experiencing during treatment, and following treatment, received verbal and written instructions on how to conduct treatment maintenance using in-vivo self-exposure in the vicinity of their own home (Öst, 2012).

Assessment meetings occurred again 1-week following treatment (post-measurement) and at follow-up (3-months and 12-months) post-treatment. These assessments were carried out by a trained therapist blinded to treatment allocation. In all cases, the BAT was completed at the end of the assessment session following self-report questionnaires. At the pre- and post-measurement occasions participants were offered the opportunity to participate in a parallel EEG study evaluating the effects of exposure therapy on neurological processes. Participants unable to visit Stockholm University at post-assessment or follow-up were offered the opportunity to complete self-report assessments by mail or online and diagnostic interviews by telephone. When assessments were conducted online or by mail, it was not possible to administer the BAT. Previous research has suggested that electronic administration of questionnaires is not different than those completed by paper (Muehlhausen et al., 2015).

2.5. Statistical analyses

Data was analyzed according to protocol (Miloff et al., 2016) on an intention-to-treat basis and supplemented by an observed-only data analysis when testing non-inferiority. Non-inferiority was evaluated only for the primary outcome variable, whereas all outcomes received superiority analysis. Non-inferiority was calculated using the double-sided confidence interval approach, by modeling post-treatment BAT scores as a function of group allocation and pre-treatment scores. Pre-treatment BAT scores were controlled for in the non-inferiority analysis to statistically adjust for a slight non-significant difference between groups prior to treatment. VRET was considered non-inferior to OST if the lower-bound of the 95% confidence interval BAT difference score following treatment was within the preset non-inferiority margin (Piaggio et al., 2012). The upper-boundary was used to establish if the novel treatment was significantly less effective than reference treatment (i.e., bound does not cross null). Similar to other studies, in the absence of prior established non-inferiority trials for specific phobia, the non-inferiority margin was selected according to an estimate of minimum change needed to determine clinically significant improvement (Andersson et al., 2013) and set at 2-points based on pre-existing data (Andersson et al., 2009). This study was powered to detect the non-inferiority margin with a standard deviation of four points (80% power). If there was truly no difference between treatments, then 100 patients would be required to be 80% sure that the lower bound of a one-sided 95% confidence interval was above the non-inferiority limit of -2 points (sealedenvelope.com).

Groups were compared on all continuous outcome measures by investigating time*group effects in full-information, maximum-likelihood mixed effects models making this a full intention-to-treat analysis (Hesser, 2015). Missing values were accounted for under the assumption of missing at random (MAR). Analysis was conducted using SPSS Version 24. Linear effects were completed in two separate models. In the first, pre to post-measure results were modeled with random intercepts (not slopes) and a variance components structure. In the second, post-measure to 12-month follow-up data was modeled using random intercepts, slopes and an unstructured covariance. Last-observation-carried-forward (LOCF) was used to conservatively account for the three participants who dropped out. Parameter estimates from simple regression models were used to estimate BAT data at post-, 3- and 12-month follow-up when SPQ scores were available from mailed or digital forms ($rs > .542$, $ps < .001$). In the intention-to-treat non-inferiority analyses, all remaining missing data were replaced with estimated values from the mixed model.

Clinically significant improvement requires that participant outcome measures at post-treatment be outside the range of a patient population (cut-off type A, Jacobson & Truax, 1991; defined as mean $\pm$ 2 SD) or within the range of a normal population if such data exists (cut-off type B or C). In the absence of normative population data for self-report measures, evaluation of clinically significant change and deterioration used cut-off type A, incorporating a 95% classification confidence safety margin (Hageman & Arrindell, 1999). Reliability figures were based on FSQ and SPQ test-retest coefficients (Muris & Merckelbach, 1996). Group-level clinically significant self-report results were computed as a conservative summation of individual-level analysis outcomes. Similar to previous studies (Andersson et al., 2009), clinically significant improvement for behavioral approach was based on pre-existing data (Öst et al., 1997) and defined as a post-treatment cut-off score of 10 or more and a minimum 2-points improvement. High end-state functioning required that the participant meet clinically significant improvement for BAT and both fear of spider self-report measures at post-treatment.

The impact of the experience of presence during VRET on primary and secondary outcome measures at post-assessment was conducted using a linear mixed model using median presence scores (split high and low), random intercepts (not slopes) and a variance components structure.

3. Results

3.1. Attrition and missing data

All participants were present at the pre-measure session, however 2 participants (1 OST; 1 VRET) dropped out prior to treatment, and 1 participant (VRET) dropped out prior to post-assessment. Three other participants at post-treatment (1 OST; 2 VRET) were provided assessments by mail meaning BAT data could not be collected.

At 3-month follow up, 88% (44/50) OST and 90% (45/50) VRET participants completed assessments. Twelve participants (27%) from each group conducted these assessments online and like mailed assessments no BAT information could be collected. At 12 months follow-up, assessments were completed by 88% (44/50) OST and 94% (47/50) VRET participants. Of these, 34% (15/44) OST participant assessments were conducted online and 9% (4/44) by mail, and in the VRET group 36% (17/47) assessments were conducted online and an additional 6% (3/47) by mail. Supportive of a missing at random assumption, no significant difference was identified in the proportion of missing data between VRET and OST groups at any time ($ps > .67$).

Participants who were present at Stockholm University for assessments were compared to participants who were not, either as a result of missing data or an assessment conducted by mail or online. Multiple tests of significance using Bonferroni corrections noted no difference on demographic and baseline outcome measures.

Table 2
Means, standard deviations and within-group effect sizes for observed data by treatment group for primary, secondary and other outcome measure at post- and follow-up periods.

	One-Session Treatment			Virtual Reality Exposure Therapy		
Time	N	Mean (SD)	Effect Size	N	Mean (SD)	Effect Size
Behavioral Approach Test (BAT)						
Pre	50	5.66 (2.47)		50	4.76 (2.71)	
Post	47	10.70 (1.68)	2.39	46	8.50 (2.29)	1.49
3 Months	32	10.91 (1.25)	2.68	34	9.03 (2.5)	1.64
12 Months	26	10.58 (1.81)	2.27	28	9.36 (1.77)	2.01
Fear of Spider Questionnaire (FSQ)						
Pre	50	98.08 (13.91)		50	95.82 (15.26)	
Post	49	54.59 (17.82)	2.72	48	70.35 (22.45)	1.33
3 Months	44	53.68 (22.34)	2.39	46	65.22 (23.76)	1.53
12 Months	44	52.2 (23.46)	2.38	48	61.94 (27.11)	1.54
Spider Phobia Questionnaire (SPQ)						
Pre	50	22.76 (2.95)		50	22.36 (3.78)	
Post	49	13.39 (4.92)	2.31	48	16.56 (5.84)	1.18
3 Months	44	13.75 (5.5)	2.04	46	15.46 (6.39)	1.32
12 Months	44	13.68 (4.73)	2.30	48	14.48 (6.4)	1.50
Patient Health Questionnaire (PHQ)						
Pre	50	2.80 (3.87)		50	2.32 (3.21)	
Post	49	2.78 (3.3)	0.01	48	1.83 (1.84)	0.19
3 Months	44	2.52 (3.15)	0.08	46	2.00 (2.71)	0.11
12 Months	44	2.75 (2.65)	0.02	48	2.38 (3.42)	0.02
Generalized Anxiety Disorder Assessment (GAD)						
Pre	50	2.90 (3.31)		50	2.32 (2.83)	
Post	49	3.02 (3.72)	0.03	48	1.98 (2.29)	0.13
3 Months	44	2.30 (3.55)	0.18	46	1.67 (2.16)	0.26
12 Months	44	2.50 (2.46)	0.14	48	2.52 (3.3)	0.07
Brunnsviken Brief Quality of Life Inventory (BBQ)						
Pre	50	74.24 (16.31)		50	77.72 (12.99)	
Post	49	75.78 (14.19)	0.10	48	78.92 (13.24)	0.09
3 Months	44	76.50 (14.37)	0.15	46	76.00 (15.78)	0.12
12 Months	44	75.34 (15.38)	0.07	48	74.92 (18.09)	0.18

Note. Effect Size = Cohen's *d* effect size.

3.2. Pre-treatment outcome measure group differences

Primary, secondary and other outcome measure observed-only mean and standard deviation results for each study assessment period are included in Table 2 alongside within-group Cohen's *d* effect sizes. A moderate non-significant pre-treatment group difference in behavioral outcome measure was observed with the VRET group scoring nearly 1 point lower on the 13-point BAT scale (indicating more fear) than the OST group, ($\beta = 0.9$, 95% CI -1.85 to 0.05, $p = .062$, $d = 0.35$). This was equivalent to a 0.9 m difference in approach distance between groups on the standardized BAT measurement. No significant difference in fear of spider self-report (FSQ, SPQ; *ps* > .45) and other outcome measures (GAD-7, PHQ-9, BBQ; *ps* > .24) were noted between the two groups at pre-treatment.

3.3. Primary and secondary outcome measures

3.3.1. Non-inferiority

Non-inferiority tests were conducted using the interval method (Piaggio et al., 2012) on intention-to-treat and observed-only BAT data at post-measure and follow-up, controlling for pre-treatment difference scores (see Table 3). VRET was not non-inferior to the OST group at post-treatment, as indicated by the lower-bound of intention-to-treat and observed data 95% confidence intervals crossing the predetermined 2-point non-inferiority margin (CI's −2.599 & −2.579, respectively; see Fig. 2). At 3-months follow-up, however, non-inferiority was obtained in the intention-to-treat analysis (CI -1.970 to −0.440) and at 12-months follow-up using both observed-only (CI -1.814 to 0.014) and intention-to-treat data (CI -1.504 to −0.111). Only the observed-only 12-month follow-up analysis, however, identified an upper-bound of the 95% CI above null (0.014). This indicates that while VRET was not inferior to OST on the primary outcome measure at 3- and 12-months, it was also significantly less effective until 12-months follow-up.

Table 3
Results of observed-only and intention-to-treat non-inferiority behavioral approach test data analysis for participants at post-, 3- and 12-month follow-up periods.

Time	Observed 95% CI			Intention-to-Treat 95% CI		
	LB	Mean	UB	LB	Mean	UB
Post	−2.579	−1.838	−1.096	−2.599	−1.819	−1.040
3 Months	−2.299	−1.365	−0.431	−1.970	−1.205	−0.440
12 Months	−1.814	−0.900	0.014	−1.504	−0.808	−0.111

Note. LB = lower bound; UB = upper bound.

3.3.2. Behavioral approach test

At 3- and 12-months follow-up, observed-only data was available for 64% (32/50) and 52% (26/50) of the OST group and 68% (34/50)

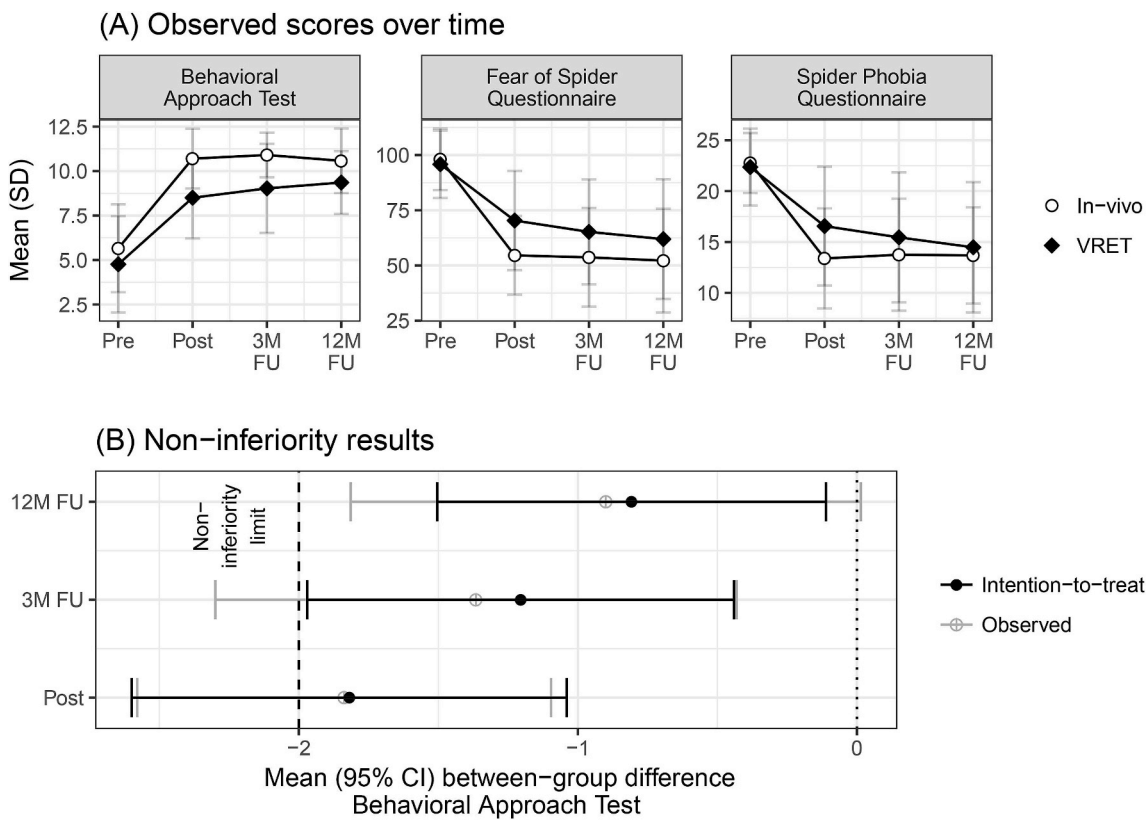

Fig. 2. (A) Mean primary and secondary longitudinal observed-only outcome measure data by treatment group at pre-, post- and follow-up assessment periods using standard deviation error bars. (B) Plotted non-inferiority analysis mean between-group difference 95% confidence intervals at post-treatment and follow-up periods using observed-only and intention-to-treat data. In-vivo = In-vivo one-session treatment; VRET = Virtual reality exposure therapy.

and 56% (28/50) of the VRET group, respectively. Fig. 2 displays longitudinal observed-only data for primary and secondary outcome measures. Linear mixed effect models noted large significant mean improvements on the behavioral avoidance measure at post-assessment for both OST (β = 4.84, 95% CI 4.15 to 5.52, $d = 2.39$) and VRET groups (β = 3.55, 95% CI 2.87 to 4.23, $d = 1.49$), however the VRET group experienced significantly less improvement than OST (β = −1.27, 95% CI -2.27 to −0.28, $p = .013$). The OST group maintained BAT treatment benefits between post and 12-months follow-up (β = −0.20, 95% CI -0.46 to 0.54, $p = .376$) whereas those treated by VRET continued to experience significant gains as compared to OST over this period (β = 0.38, 95% CI .13 to .63, $p = .002$).

3.3.3. Self-report (FSQ and SPQ)

Strong reductions in self-reported fear of spiders (FSQ and SPQ) were found for both VRET and OST at post-treatment. FSQ scores in the OST group were reduced by 42.76 points (95% CI -48.69 to −36.83, $d = 2.72$) and in the VRET group by 24.38 points (95% CI -30.31 to −18.46, $d = 1.33$). SPQ measures in the OST group were reduced by 9.12 points (95% CI -10.58 to −7.66, $d = 2.31$) and in the VRET group by 5.52 points (95% CI -6.98 to −4.06, $d = 1.18$). At post-treatment the VRET group was significantly less improved on both the FSQ (β = 18.38, 95% CI 10.00 to 26.76, $p < .001$) and SPQ self-report outcome measures than the OST group (β = 3.60, 95% CI 1.54 to 5.66, $p = .001$). Linear mixed effect model analysis of FSQ self-report outcomes between post- and 12-month follow-up indicate that treatment benefits were maintained in both OST (β = −1.33, 95% CI -4.46 to 1.80, $p = .40$) and VRET groups (β = −4.04, 95% CI -7.07 to −1.01, $p = .22$) with no significant increase in fear symptoms. Data indicate that SPQ self-report results over 12-month follow-up were maintained for the OST group (β = 0.19, 95% CI -0.49 to 0.87, $p = .58$) but continued to improve significantly in the VRET group (β = −1.00, 95% CI -1.66 to −0.34, $p = .014$).

3.3.4. Diagnostic criteria

Participants meeting diagnostic criteria for fear of spiders at 3- and 12-months follow-up were not significantly different between groups with 13% (5/38) OST and 31% (11/36) VRET participants meeting criteria at the earlier time period, X^2 (1, $N = 74$) = 3.302, $p = .069$, and just 3 participants (2/28 VRET; 1/31 OST) meeting criteria at the later. A survival analysis was therefore conducted assuming criteria was met for all participants unable or unwilling to complete a diagnostic interview. Chi-square tests performed on diagnostic criteria survival data indicated that at both 3- and 12-months follow-up no significant difference was identified between OST and VRET ($ps > .105$).

Therapists recorded whether they were unblinded during the course of the 12-month follow-up diagnostic interview. In total, therapists were unblinded in 11% (6/57) of interviews with no significant difference between groups.

3.4. Other analyses

3.4.1. Clinically significant change and deterioration

Clinically significant improvement according to the BAT was identified in 72% (36/50) OST and 28% (14/50) of VRET participants at post-measure, X^2 (1, $N = 100$) = 19.360, $p < .001$. Both FSQ and SPQ self-report noted clinically significant improvement in 56% (28/50) of OST and 38% (19/50) of VRET participants, with no significant difference between them, X^2 (1, $N = 100$) = 3.252, $ps > .071$. High end-state functioning was identified for a significantly greater number of OST participants (22/50) than VRET participants (11/50) at post-treatment, X^2 (1, $N = 100$) = 5.473, $p = .019$. No participant was identified as having reliable deterioration according to fear of spider self-report measures and just 1 participant (VRET group) had worse

BAT outcomes at post- as compared to pre-treatment.

3.4.2. Other outcome measures (PHQ-9, GAD-7, BBQ)

No significant difference was identified between VRET and OST on other outcome measures including BBQ, PHQ-9 or GAD-7 using linear mixed models for between pre- and post-treatment, as well as post-treatment and follow-up periods ($ps > .50$).

3.4.3. Negative effects questionnaire (NEQ-32)

Mean OST and VRET scores on the NEQ-32, signifying intensity and impact of negative effects, were M = 3.18 (SD = 4.23) and M = 4.74 (SD = 6.20), respectively and not significantly different at post-treatment ($p = .145$). The three most highly reported negative effects in the OST group were "I felt more worried" (13/48), followed by "I experienced more anxiety" (12/48) and "I experienced more unpleasant feelings" (11/48), whereas the most highly reported negative effects of VRET participants were "I felt that my expectations for the treatment were not fulfilled" (19/47), followed by "I felt that the treatment did not produce any results" (18/47) and "I felt that the treatment did not suit me" (15/47). The proportion of participants with any occurrence of negative effect was not significantly different between groups, X^2 (1, $N = 100$) = 0.041, $p = .84$.

3.4.4. BAT maximum score (12 points)

Completion of the BAT with a score of 12 points, the maximum possible on the scale, was accomplished by a greater of number of participants in the OST group than the VRET group at post-, 3- and 12-months follow-up ($ps < .016$). Approx. 50% of participants in the OST group achieved a score of 12-points at post-, 3- and 12-months (25/47, 12/26, 16/32) vs. 11–15% in the VRET group over the same time periods (5/46, 5/34, 4/28), respectively.

3.4.5. Impact of presence

No significant difference in BAT scores at post-treatment were observed for those with high-presence as compared to those with low-presence (β = 0.90, 95% CI -0.32 to 2.13, $p = .14$). Similar non-significant findings for the effect of presence on self-reported fear of spiders (FSQ and SPQ) were observed at post-assessment ($ps > 0.16$).

4. Discussion

Strong reductions in spider phobia symptom and behavioral avoidance were noted in both VRET and OST groups in this study. At post-assessment, 1-week following treatment, in-vivo exposure therapy was superior to the virtual reality treatment and maintained its benefits over follow-up. The VRET group, however, caught up with OST. Non-inferiority was achieved at 3- and 12-months follow-up although significantly worse until 12-months. The mean difference in BAT scores post-treatment of less than 2-points amounted to what may be considered over-learning (reaching beyond normal non-phobic behavior), the difference between physically touching a spider and holding one's hand near it. No deterioration was noted in either treatment group and there was no significant difference on the negative treatment effects questionnaire. As hypothesized, results suggest that a single-session of VRET is not worse than OST in regards to the primary outcome measure (BAT), however non-inferiority was identified not until 3-months follow-up rather than immediately after treatment, as was expected.

An important difference here is that few VRET research studies have been conducted using an automated treatment not requiring engagement of a therapist (Freeman et al., 2017). Virtual reality studies have tended to replace in-vivo with virtual stimuli but the human therapist continued to be an essential requirement for treatment to be conducted. To date, only one automated VR treatment study has been published for fear of heights (Freeman et al., 2018) and one research protocol (Donker, Van Esveld, Fischer, & Van Straten, 2018). Freeman et al. (2018) identified large (Cohen's $d = 2.00$) effect size reductions in self-reported fear as compared to wait-list at post- and 1-month follow-up in a clinical population. In this study, as in Freeman et al. (2018), a therapist was present with the participant for safety reasons and fitting of the headset, with therapist-level support limited to a small minority of participants. Otherwise the VIMSE application (which constituted the treatment) was entirely self-contained. The feasibility of conducting automated VR exposure treatments at home (rather than a laboratory setting) will be tested by Donker et al. (2018) who plan to mail inexpensive cardboard-based goggles directly to participants. This format is likely to be effective; a recent VRET study for fear of public speaking found large within-group effect size reductions of self-reported anxiety using similar mailed goggles and self-led psychoeducational material ($d = 1.38$; Lindner et al., 2018). Compared to specific phobias such as fear of flying, treatment of spider phobia is relatively inexpensive given the ease of accessing stimuli during certain times of the year (Choy et al., 2007), however psychological services continue to be expensive for individuals and society (Andlin-Sobocki & Wittchen, 2005).

At post-treatment, effect sizes in the VRET group for behavioral avoidance ($d = 1.49$) were similar to past within-group design spider phobia studies including Bouchard et al. (2006; $d = 1.46$), and Côté and Bouchard (2005; $d = 1.48$) as well as earlier studies from Garcia-Palacios et al. (2002; $d = 1.24$) and Hoffman et al. (2003; $d = 1.24$) who measured avoidance using distance travelled (meters and feet, respectively) rather than a demarcated scale. Reductions in self-reported fear of spiders (FSQ), however, were smaller ($d = 1.33$) as compared to previous VRET research at post-treatment such as Côté and Bouchard (2005; $d = 2.68$), and Bouchard et al. (2006; $d = 3.15$). In Michaliszyn et al. (2010), the only other randomized controlled VRET trial to date for spider phobia in adults with an in-vivo exposure therapy control group, larger within-group BAT improvements ($d = 2.03$) and self-reported fear improvements (FSQ; $d = 2.91$) were obtained. It should be noted that in Michaliszyn et al. (2010), VRET treatments were conducted by a human therapist, participants (n = 3) were moved to an in-vivo exposure group if they failed to react to virtual spiders, and a greater number of VRET sessions were used (8 sessions x 1.5 h).

This study adds to the limited evidence that a single VRET session can provide significant benefit for spider phobia sufferers. Using a single-session format consisting of six 5 min stereoscopic 3D video loops displayed on an Oculus Rift headset, the most recent spider phobia VRET study published relied on real spider footage rather than animated computer graphics (Minns et al., 2018). The authors noted that real footage rather than simulations can be cheaper to produce and more realistic; however, efficacy was lower, with medium BAT within-group effect sizes (demarcated scale; $d = 0.47$) and smaller improvements than previous studies on self-reported fear (FSQ; $d = 0.85$). It is worth noting that participants were not required to meet diagnostic criteria in this study, treatment time was much shorter than previous studies, and post-assessment was conducted immediately following treatment rather than 1-week later. Another study evaluating the impact of short (45 min) single-sessions of VRET on spider phobia patients (with or without post-treatment sleep) also indicated significant benefits for participants, particularly those in the sleep condition (Kleim et al., 2014); however, an alternative measure of self-reported fear and behavioral avoidance was used making comparison difficult. Until now only augmented reality (AR), a related technology in which computer animated stimuli appear to be projected onto real world surfaces, has been compared against state-of-the-art OST in a single-session format for the treatment of small animal phobias (spider and cockroach; Botella et al., 2016). In the study, both OST and AR sessions lasted up to 3 h. Similarly, OST was more effective at post-assessment but participants continued to improve at follow-up in the AR condition, with no significant difference noted at 3 and 6-months.

A number of limitations of this study must be noted. For one, loss to follow-up was minimized by allowing participants in some cases to conduct assessments online or in writing. This ensured low study attrition but meant that not all participants were able to complete the

primary outcome measure. Therapists were trained to ensure proficiency with OST and participated in supervision with an experienced psychologist; however, their skills were not evaluated systematically prior to performing the intervention. Treatment sessions were not video-taped for compliance checks against manualized OST. Therefore, although effect sizes on the primary outcome measure in this study are similar to previous OST trials ($d = 2.39$ vs. 2.73; Andersson et al., 2009), the precision of the OST conducted in this study may be closer to that of an effectiveness trial performed in a clinical setting. Future studies must protect against "biocreep," in which the next generation of non-inferiority trial uses a slightly inferior reference treatment ad infinitum until the active control is no better than placebo (D'Agostino, Massaro & Sullivan, 2003). Although, the VRET treatment was an automated process and therapists were instructed to act as a "computer technician," therapist-level support was provided in a number of cases and non-specific therapeutic factors may have accrued from their presence. It is also not clear the degree to which post-treatment maintenance using self-exposure, which was recommended to participants of the trial, or very short exposures to in-vivo spiders during BAT assessments, had an effect on long-term outcomes and data on this subject is unfortunately limited (Wolitzky-Taylor et al., 2008).

VRET software applications such as VIMSE are still in an early stage of development. Despite the inclusion of multiple therapeutic components such as voice-guided virtual therapist (Rizzo et al., 2011), gamification elements (Johnson et al., 2016) and inhibitory learning strategies (Craske, Treanor, Conway, Zbozinek, & Vervliet, 2014) using exposure scenarios involving multiple contexts (Shiban, Pauli, & Mühlberger, 2013) and multiple spider stimuli (Shiban, Schelhorn, Pauli, & Mühlberger, 2015), in addition to game mechanics and educational content intended to create a positive valence towards spiders (Dour, Brown, & Craske, 2016), dismantling studies will be needed to determine which characteristics imparted treatment benefits.

Treatments for specific phobia using VRET, have for more than 20 years provided evidence of positive treatment effects for spider phobia and other specific phobias. Until recently, however, this technology was inaccessible to all but a few at well-funded universities and specialized clinics limiting dissemination (Freeman et al., 2017). This study is the first randomized non-inferiority trial to evaluate VRET, confirming that the technology is effective at reducing fear even with a technician-assisted low-cost consumer-available device and limited therapist support. While evidence for the benefits of automated VRET treatments are increasing, they are still not ready to replace a therapist for providing high-quality assessment, assisting patients on identifying a correct course of treatment, and determining treatment completion. Prior to widespread dissemination, effectiveness trials will ensure the feasibility of conducting VRET outside the laboratory, such as in the home as a self-help treatment or in other real-world settings such as hospitals and private practice (Botella, Fernández-Álvarez, Guillén, García-Palacios, & Baños, 2017a). The negative concerns of practicing therapists towards VRET will need to be addressed. A recent cross-sectional survey of cognitive-behavioral therapists found significantly more positive than negative attitudes ($p < .001$); however, average negative attitude was a stronger negative predictor of self-rated likelihood of future use than positive attitude was a positive predictor (Linder et al., 2019). Preliminary evidence from the negative treatment effects questionnaire suggest that VRET may be experienced as less aversive to participants than OST, reducing barriers to its use as a self-help treatment. The experience of presence in the VR environment, although highly correlated with self-reported anxiety in previous trials of animal phobias (Ling, Nefs, Morina, Heynderickx, & Brinkman, 2014), did not appear to constitute a requirement for treatment efficacy in this study.

Evidence of non-inferiority at 3- and 12-months follow-up, although significantly less effective until 12-months, nevertheless demonstrates that while in-vivo exposure therapy remains the gold-standard treatment for specific phobia, the benefits of VRET are not worse over follow-up. This research adds support for the efficacy of virtual reality-based treatments in general, whose benefits recent reviews of the field have indicated extend beyond exposure therapy (Botella, Baños, García-Palacios, & Quero, 2017b).

Declaration of interest

WH is the founder of the private company that developed the VIMSE application (Mimerse AB). The application was modified and released for commercial sale as ITSY-Spider Phobia Treatment Game for VR on the Samsung Gear VR store (8 March 2016). WH was not involved in the data analysis or any decision related to publication of this article. PL has received consulting fees from the same company but has no financial stake and was not primarily responsible for analysis. The other authors declare that they have no competing interests.

Acknowledgements

The authors would like to thank Katherine (Katie) Cotter for her many contributions including data collection, Ove Sigfridsson from Stockholm Vatten och Avfall for providing ready access to spiders during the cold winter months of the study, Anna-Clara Hellstadius for space to store these spiders, Oskar Henrikson for the voice-over acting as the virtual therapist in the Swedish version of VIMSE, and the multiple real-life therapists who provided treatment and support for participants in this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.brat.2019.04.004.

Funding

This research was supported by a grant from the Swedish Government innovation agency Vinnova. The grant was part of the "Forska och Väx"-programme awarded to Mimerse AB. The grant was mainly used to fund development of the VRET software.

References

American Psychological Association (2013). *Diagnostic and statistical manual of mental disorders (Fifth). Washington, DC*.

Andersson, G., Hesser, H., Veilord, A., Svedling, L., Andersson, F., Sleman, O., et al. (2013). Randomised controlled non-inferiority trial with 3-year follow-up of internet-delivered versus face-to-face group cognitive behavioural therapy for depression. *Journal of Affective Disorders, 151*(3), 986–994. https://doi.org/10.1016/j.jad.2013.08.022.

Andersson, G., Waara, J., Jonsson, U., Malmaeus, F., Carlbring, P., & Öst, L. (2009). Internet based self help versus one session exposure in the treatment of spider phobia: A randomized controlled trial. *Cognitive Behaviour Therapy, 38*(2), 114–120. https://doi.org/10.1080/16506070902931326.

Andlin-Sobocki, P., & Wittchen, H. U. (2005). Cost of anxiety disorders in Europe. *European Journal of Neurology, 12*(SUPPL. 1), 39–44. https://doi.org/10.1111/j.1468-1331.2005.01196.x.

Barlow, D. H. (2002). *Anxiety and its disorders: The nature and treatment of anxiety and panic, American Journal of psychiatry* (2nd ed.). New York: The Guilford Press. https://doi.org/10.1176/appi.ajp.159.8.1453.

Botella, C., Baños, R. M., García-Palacios, A., & Quero, S. (2017b). Virtual reality and other realities. *The science of cognitive behavioral therapy* (pp. 551–590). . https://doi.org/10.1016/B978-0-12-803457-6.00022-2.

Botella, C., Breton-López, J., Quero, S., Baños, R. M., García-Palacios, A., Zaragoza, I., et al. (2011). Treating cockroach phobia using a serious game on a mobile phone and augmented reality exposure: A single case study. *Computers in Human Behavior, 27*(1), 217–227. https://doi.org/10.1016/j.chb.2010.07.043.

Botella, C., Fernández-Álvarez, J., Guillén, V., García-Palacios, A., & Baños, R. (2017a). Recent progress in virtual reality exposure therapy for phobias: A systematic review. *Current Psychiatry Reports, 19*(7), 42. https://doi.org/10.1007/s11920-017-0788-4.

Botella, C., Pérez-Ara, M.Á., Bretón-López, J., Quero, S., García-Palacios, A., & Baños, R. M. (2016). In vivo versus augmented reality exposure in the treatment of small animal phobia: A randomized controlled trial. *PLoS One, 11*(2), 1–22. https://doi.org/10.1371/journal.pone.0148237.

Bouchard, S., Côté, S., St-Jacques, J., Robillard, G., & Renaud, P. (2006). Effectiveness of virtual reality exposure in the treatment of arachnophobia using 3D games.

Technology and Health Care, 14, 19–27.
Carl, E., Stein, A. T., Levihn-Coon, A., Pogue, J. R., Rothbaum, B., Emmelkamp, P., Asmundson, G. J. G., Carlbring, P., & Powers, M. B. (2019). Virtual reality exposure therapy for anxiety and related disorders: A meta-analysis of randomized controlled trials. *Journal of Anxiety Disorders, 61*, 27–36. https://doi.org/10.1016/j.janxdis.2018.08.003.
Carlin, A. S., Hoffman, H. G., & Weghorst, S. (1997). Virtual reality and tactile augmentation in the treatment of spider phobia : A case report. *Behaviour Research and Therapy, 35*(2), 153–158.
Chan, A., Tetzlaff, J., Altman, D., Laupacis, A., Gøtzsche, P., & Krleža-Jerić, K. (2013). Research and reporting methods annals of internal medicine SPIRIT 2013 Statement : Defining standard protocol items for clinical trials. *Annals of Internal Medicine, 158*(3), 200–207. https://doi.org/10.7326/0003-4819-158-3-201302050-00583.
Choy, Y., Fyer, A. J., & Lipsitz, J. D. (2007). Treatment of specific phobia in adults. *Clinical Psychology Review,* (27), 266–286. https://doi.org/10.1016/j.cpr.2006.10.002.
Côté, S., & Bouchard, S. (2005). Documenting the efficacy of virtual reality exposure with psychophysiological and information processing measures. *Applied Psychophysiology Biofeedback, 30*(3), 217–232. https://doi.org/10.1007/s10484-005-6379-x.
Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy, 58*, 10–23. https://doi.org/10.1016/j.brat.2014.04.006.
D'Agostino Sr, R. B., Massaro, J. M., & Sullivan, L. M. (2003). Non-inferiority trials: Design concepts and issues–the encounters of academic consultants in statistics. *Statistics in Medicine, 22*(2), 169–186. https://doi.org/10.1002/sim.1425.
Deacon, B. J., Farrell, N. R., Kemp, J. J., Dixon, L. J., Sy, J. T., Zhang, A. R., et al. (2013). Assessing therapist reservations about exposure therapy for anxiety disorders: The Therapist Beliefs about Exposure Scale. *Journal of Anxiety Disorders, 27*(8), 772–780. http://doi.org/10.1016/j.janxdis.2013.04.006.
Donker, T., Van Esveld, S., Fischer, N., & Van Straten, A. (2018). 0Phobia–towards a virtual cure for acrophobia: Study protocol for a randomized controlled trial. *Trials, 19*(1), 433. https://doi.org/10.1186/s13063-018-2704-6.
Dour, H. J., Brown, L. A., & Craske, M. G. (2016). Positive valence reduces susceptibility to return of fear and enhances approach behavior. *Journal of Behavior Therapy and Experimental Psychiatry, 50*, 277–282. https://doi.org/10.1016/j.jbtep.2015.09.010.
Fernández-Álvarez, J., Rozental, A., Carlbring, P., Colombo, D., Riva, G., Anderson, P. L., et al. (2018). Deterioration rates in Virtual Reality Therapy: An individual patient data level meta-analysis. *Journal of Anxiety Disorders, 61*, 3–17. https://doi.org/10.1016/j.janxdis.2018.06.005.
First, M. B., Robert, L., Gibbon, M., & Williams, J. B. W. (2002). *Structure clinical interview for DSM-IV-TR Axis I, research version.* New York: Biometric Research, New York State Psychiatric Institute Patient Edition (SCID-I/P).
Freeman, D., Haselton, P., Freeman, J., Spanlang, B., Kishore, S., Albery, E., et al. (2018). Automated psychological therapy using immersive virtual reality for treatment of fear of heights: A single-blind, parallel-group, randomised controlled trial. *The Lancet Psychiatry, 5*(8), 625–632. https://doi.org/10.1016/S2215-0366(18)30226-8.
Freeman, D., Reeve, S., Robinson, A., Ehlers, A., Clark, D., Spanlang, B., et al. (2017). Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychological Medicine, 1–8*https://doi.org/10.1017/S003329171700040X.
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 & Behavior : The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society, 10*(5), 722–724. https://doi.org/10.1089/cpb.2007.9962.
Garcia-Palacios, A., Hoffman, H., Carlin, A., Furness, T. A., & Botella, C. (2002). Virtual reality in the treatment of spider phobia: A controlled study. *Behaviour Research and Therapy, 40*, 983–993. https://doi.org/10.1016/S0005-7967(01)00068-7.
Hageman, W. J., & Arrindell, W. A. (1999). Establishing clinically significant change: increment of precision and the distinction between individual and group level of analysis. *Behaviour Research and Therapy, 37*, 1169–1193.
Hesser, H. (2015). Modeling individual differences in randomized experiments using growth models: Recommendations for design, statistical analysis and reporting of results of internet interventions. *Internet Interventions,* (2), 110–120. https://doi.org/10.1016/j.invent.2015.02.003.
Hoffman, H. G., Garcia-Palacios, A., Carlin, A., Furness, T. A., & Botella-Arbona, C. (2003). Interfaces that heal: Coupling real and virtual objects to treat spider phobia. *International Journal of Human-Computer Interaction, 16*(2), 283–300.
Jacobson, N. S., & Truax, P. (1991). Clinical Significance : A statistical approach to defining meaningful change in psychotherapy research. *Journal Consulting and Clinical Psychology, 59*(1), 12–19.
Johnson, D., Deterding, S., Kuhn, K. A., Staneva, A., Stoyanov, S., & Hides, L. (2016). Gamification for health and wellbeing: A systematic review of the literature. *Internet Interventions, 6*, 89–106. https://doi.org/10.1016/j.invent.2016.10.002.
Kleim, B., Wilhelm, F. H., Temp, L., Margraf, J., Wiederhold, B. K., & Rasch, B. (2014). Sleep enhances exposure therapy. *Psychological Medicine, 44*(7), 1511–1519. https://doi.org/10.1017/S0033291713001748.
Klorman, R., Weerts, T. C., Hastings, J. E., Melamed, B. G., & Lang, P. J. (1974). Psychometric description of some specific-fear questionnaires. *Behavior Therapy, 5*, 401–409. https://doi.org/10.1016/S0005-7894(74)80008-0.
Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9 : A new depression measure. *Psychiatric Annals, 32*(9), 509–515. https://doi.org/170553651.
Lieb, R., Miche, M., Gloster, A. T., Beesdo-Baum, K., Meyer, A. H., & Wittchen, H. U. (2016). Impact of specific phobia on the risk of onset of mental disorders: A 10-year prospective-longitudinal community study of adolescents and young adults. *Depression and Anxiety, 33*(7), 667–675. http://doi.org/10.1002/da.22487.
Lindner, P., Frykheden, O., Forsström, D., Andersson, E., Ljótsson, B., Hedman, E., et al. (2016). The brunnsviken brief quality of life scale (BBQ): Development and psychometric evaluation. *Cognitive Behaviour Therapy, 6073*(March), 1–14. https://doi.org/10.1080/16506073.2016.1143526.
Lindner, P., Miloff, A., Fagernäs, S., Andersen, J., Sigeman, M., Andersson, G., et al. (2018). 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. https://doi.org/10.1016/j.janxdis.2018.07.003.
Lindner, P., Miloff, A., Hamilton, W., Reuterskiöld, L., Andersson, G., Powers, M. B., 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, 6073*(March), 1–17. https://doi.org/10.1080/16506073.2017.1280843.
Lindner, P., Miloff, A., Zetterlund, E., Reuterskiöld, L., Andersson, G., & Carlbring, P. (2019). Attitudes towards 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*. https://doi.org/10.3389/fpsyg.2019.00176.
Lindner, P., Nyström, M. B. T., Hassmén, P., Andersson, G., & Carlbring, P. (2015). Who seeks ICBT for depression and how do they get there? Effects of recruitment source on patient demographics and clinical characteristics. *Internet Interventions, 2*(2), 221–225. https://doi.org/10.1016/j.invent.2015.04.002.
Ling, Y., Nefs, H. T., Morina, N., Heynderickx, I., & Brinkman, W. P. (2014). A meta-analysis on the relationship between self-reported presence and anxiety in virtual reality exposure therapy for anxiety disorders. *PLoS One, 9*(5), 1–12. https://doi.org/10.1371/journal.pone.0096144.
Lobbestael, J., Leurgans, M., & Arntz, A. (2011). Inter-rater reliability of the structured clinical interview for DSM-IV Axis I disorders (SCID I) and Axis II disorders (SCID II). *Clinical Psychology and Psychotherapy, 18*(18), 75–79. https://doi.org/10.1002/cpp.
McCann, R. A., Armstrong, C. M., Skopp, N. A., Edwards-Stewart, A., Smolenski, D. J., June, J. D., et al. (2014). Virtual reality exposure therapy for the treatment of anxiety disorders: An evaluation of research quality. *Journal of Anxiety Disorders, 28*(6), 625–631. https://doi.org/10.1016/j.janxdis.2014.05.010.
Michaliszyn, D., Marchand, A., Bouchard, S., Martel, M.-O., & Poirier-Bisson, J. (2010). A randomized, controlled clinical trial of in virtuo and in vivo exposure for spider phobia. *Cyberpsychology, Behavior, and Social Networking, 13*(6), 689–695.
Miloff, A., Lindner, P., Hamilton, W., Reuterskiöld, L., Andersson, G., & Carlbring, P. (2016). Single-session gamified virtual reality exposure therapy for spider phobia vs. traditional exposure therapy: Study protocol for a randomized controlled non-inferiority trial. *Trials, 17*(1), 60. https://doi.org/10.1186/s13063-016-1171-1.
Minns, S., Levihn-Coon, A., Carl, E., Smits, J. A. J., Miller, W., Howard, D., et al. (2018). Immersive 3D exposure-based treatment for spider fear: A randomized controlled trial. *Journal of Anxiety Disorders, 58*, 1–7. https://doi.org/10.1016/j.janxdis.2018.05.006.
Morina, N., Ijntema, H., Meyerbröker, K., & Emmelkamp, P. M. G. (2015). Can virtual reality exposure therapy gains be generalized to real-life? A meta-analysis of studies applying behavioral assessments. *Behaviour Research and Therapy, 74*, 18–24. https://doi.org/10.1016/j.brat.2015.08.010.
Muehlhausen, W., Doll, H., Quadri, N., Fordham, B., O'Donohoe, P., Dogar, N., et al. (2015). Equivalence of electronic and paper administration of patient-reported outcome measures: A systematic review and meta-analysis of studies conducted between 2007 and 2013. *Health and Quality of Life Outcomes, 13*, 167. https://doi.org/10.1186/s12955-015-0362-x.
Mühlberger, A., Weik, A., Pauli, P., & Wiedemann, G. (2006). One-session virtual reality exposure treatment for fear of flying: 1-year follow-up and graduation flight accompaniment effects. *Psychotherapy Research, 16*(1), 26–40. https://doi.org/10.1080/10503300500090944.
Muris, P., & Merckelbach, H. (1996). A comparison of two spider fear questionnaires. *Journal of Behavior Therapy and Experimental Psychiatry, 27*(3), 241–244. https://doi.org/10.1016/S0005-7916(96)00022-5.
Newman, M. G., Szkodny, L. E., Llera, S. J., & 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*(1), 89–103. https://doi.org/10.1016/j.cpr.2010.09.008.
Olatunji, B. O., Deacon, B. J., & Abramowitz, J. S. (2009). The cruelest cure? Ethical issues in the implementation of exposure-based treatments. *Cognitive and Behavioral Practice, 16*(2), 172–180. https://doi.org/10.1016/j.cbpra.2008.07.003.
Ollendick, T. H., Öst, L. G., Reuterskiöld, L., Costa, N., Cederlund, R., Sirbu, C., et al. (2009). One-session treatment of specific phobias in youth: A randomized clinical trial in the United States and Sweden. *Journal of Consulting and Clinical Psychology, 77*(3), 504. https://doi.org/10.1037/a0015158.
Oosterink, F., De Jongh, A., & Hoogstraten, J. (2009). Prevalence of dental fear and phobia relative to other fear and phobia subtypes. *European Journal of Oral Sciences, 117*(2), 135–143. http://doi.org/10.1111/j.1600-0722.2008.00602.x.
Opriş, D., Pintea, S., García-Palacios, A., Botella, C., Szamosközi, Ş., & David, D. (2012). Virtual reality exposure therapy in anxiety disorders: A quantitative meta-analysis. *Depression and Anxiety, 29*(37), 85–93. https://doi.org/10.1002/da.20910.
Öst, L.-G. (1989). One-Session treatment for specific phobias. *Behaviour Research and Therapy, 27*(I), 1–7.
Öst, L.-G. (2012). One-session treatment: Principles and procedures with adults. In T. E. DavisIII, T. Ollendick, & L.-G. Öst (Eds.). *Intensive one-session treatment of specific phobias* (pp. 55–95). New York: Springer. https://doi.org/10.1007/978-1-4614-3253-1_4.
Öst, L.-G., Ferebee, I., & Furmark, T. (1997). One-session group therapy of spider phobia: Direct versus indirect treatments. *Behaviour Research and Therapy, 35*(8), 721–732.
Öst, L.-G., Salkovskis, P. M., & Hellström, K. (1991). One-session therapist-directed

exposure vs. self-exposure in the treatment of spider phobia. *Behavior Therapy, 22*(3), 407–422. https://doi.org/10.1016/S0005-7894(05)80374-0.

Page, S., & Coxon, M. (2016). Virtual reality exposure therapy for anxiety disorders: Small samples and no controls? *Frontiers in Psychology, 7*(326), 1–4. https://doi.org/10.3389/fpsyg.2016.00326.

Parsons, T. D. (2015). Virtual reality exposure therapy for anxiety and specific phobias category: Psychology and human behavior when. *Encyclopedia of information science and technology* (pp. 288–296). (Third). IGI Global. https://doi.org/10.4018/978-1-4666-5888-2.ch636.

Parsons, T. D., & Rizzo, A. A. (2008). Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry, 39*, 250–261. https://doi.org/10.1016/j.jbtep.2007.07.007.

Piaggio, G., Elbourne, D. R., Altman, D. G., Pocock, S. J., & Evans, S. J. W. (2012). Reporting of noninferiority and equivalence randomized trials: An extension of the CONSORT statement. *JAMA : The Journal of the American Medical Association, 308*(24), 2594–2604. https://doi.org/10.1001/jama.2012.87802.

Rizzo, A., Lange, B., Buckwalter, J. G., Forbell, E., Kim, J., Sagae, K., et al. (2011). SimCoach : An intelligent virtual human system for providing healthcare information and support. *International Journal on Disability and Human Development, 10*(4), 277–281. https://doi.org/10.1515/IJDHD.2011.046.

Rozental, A., Kottorp, A., Boettcher, J., Andersson, G., & Carlbring, P. (2016). Negative effects of psychological treatments: An exploratory factor analysis of the negative effects questionnaire for monitoring and reporting adverse and unwanted events. *PLoS One, 11*(6), 1–22. https://doi.org/10.1371/journal.pone.0157503.

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor Analytic insights. *Presence: Teleoperators and Virtual Environments, 10*(3), 266–281. https://doi.org/10.1162/105474601300343603.

Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *BMC Medicine, 8*(1), 18. http://doi.org/10.1186/1741-7015-8-18.

Shiban, Y., Pauli, P., & Mühlberger, A. (2013). Effect of multiple context exposure on renewal in spider phobia. *Behaviour Research and Therapy, 51*(2), 68–74. https://doi.org/10.1016/j.brat.2012.10.007.

Shiban, Y., Schelhorn, I., Pauli, P., & Mühlberger, A. (2015). Effect of combined multiple contexts and multiple stimuli exposure in spider phobia: A randomized clinical trial in virtual reality. *Behaviour Research and Therapy, 71*, 45–53. https://doi.org/10.1016/j.brat.2015.05.014.

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Lo, B. (2006). A brief measure for assessing generalized anxiety disorder. *JAMA Internal Medicine, 166*(10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092.ABSTRACT.

St-Jacques, J., Bouchard, S., & Bélanger, C. (2010). Is virtual reality effective to motivate and raise interest in phobic children toward therapy? A clinical trial study of in vivo with in virtuo versus in vivo only treatment exposure. *Journal of Clinical Psychiatry, 71*(7), 924–931. https://doi.org/10.4088/JCP.08m04822blu.

Szymanski, J., & O'Donohue, W. (1995). Fear of spiders questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry, 26*(1), 31–34. https://doi.org/10.1016/0005-7916(94)00072-T.

Turner, W. A., & Casey, L. M. (2014). Outcomes associated with virtual reality in psychological interventions: Where are we now? *Clinical Psychology Review, 34*(8), 634–644. https://doi.org/10.1016/j.cpr.2014.10.003.

Valmaggia, L. R., Latif, L., Kempton, M. J., & Rus-Calafell, M. (2016). Virtual reality in the psychological treatment for mental health problems: An systematic review of recent evidence. *Psychiatry Research, 236*, 189–195. https://doi.org/10.1016/j.psychres.2016.01.015.

Vlaescu, G., Alasjö, A., Miloff, A., Carlbring, P., & Andersson, G. (2016). Features and functionality of the Iterapi platform for internet-based psychological treatment. *Internet Interventions, 6*, 107–114. https://doi.org/10.1016/j.invent.2016.09.006.

Wardenaar, K. J., Lim, C. C. W., Al-Hamzawi, A. O., Alonso, J., Andrade, L. H., Benjet, C., et al. (2017). The cross-national epidemiology of specific phobia in the World Mental Health Surveys. *Psychological Medicine, 47*(10), 1–17. https://doi.org/10.1017/S0033291717000174.

Wolitzky-Taylor, K. B., Horowitz, J. D., Powers, M. B., & Telch, M. J. (2008). Psychological approaches in the treatment of specific phobias: A meta-analysis. *Clinical Psychology Review, 28*, 1021–1037. https://doi.org/10.1016/j.cpr.2008.02.007.

Zlomke, K., & Davis, T. E., III (2008). One-session treatment of specific phobial: A detailed description and review of treatment efficacy. *Behaviour Therapy, 39*, 207–223. https://doi.org/10.1016/j.beth.2007.07.003.