Title: Systematic review and meta-analysis of randomised controlled trials for evaluating the effectiveness of virtual reality therapy for social anxiety disorder

Abstract

Objective

To evaluate the effectiveness of VR therapy (VRT) for symptoms related to social anxiety disorder (SAD), namely fear and avoidance of social interactions and performance situations (FASIP), fear of negative evaluation (FNE), anxiety and depression, a systematic review and meta-analysis were performed.

Methods

Medline, PubMed, Science Direct, Web of Science, CINAHL, PsychINFO and Scopus were searched to include randomised controlled trials of VRT for SAD that met the criteria. A total of 15 RCTs with 720 participants published between 1998 and 2022 were included. Hedge's g with a 95% confidence interval (CI) was adopted to compute the effect sizes.

Results

Results showed no difference between the effect of VRT and CBT on FASIP, FNE, anxiety and depression and a large effect size for VRT versus the waitlist control group on FASIP (g = -1.170, 95% CI: -2.056–0.283; $p < 0.010$). The moderator analysis demonstrated that VRT was superior to the controlled group in addressing FASIP, FNE and anxiety when the sample size was smaller than 50 and the number of sessions was five or fewer.

Limitations

Differences in hardware, software and intervention duration for VRT across studies

Conclusion

This study confirmed the feasibility of VRT in alleviating the FASIP in patients with SAD, with the waitlist control group as a comparison. However, the effectiveness of VRT was not significant in FASIP, FNE, anxiety and depression compared to cognitive behavioural therapy (CBT). Additional social interaction scenarios should be developed in VRT, standardised hardware should be used and the proper length of exposure time to VR should be determined to enhance the efficacy of VRT.

Keywords:

Virtual reality therapy; Social anxiety disorder; Fear of social interactions; Fear of negative evaluation; Depression; Randomised controlled trials

1. Introduction

Social anxiety disorder (SAD), also known as social phobia, is the most ubiquitous psychiatric disorder with a lifetime prevalence of 12% (Kessler et al., 2005). SAD causes extreme, consistent fear in different social settings in which individuals anticipate negative judgments from others and feel that their presence makes others uncomfortable (Gilbert, 2001). The physical symptoms associated with SAD include blushing, nausea, excessive sweating, trembling, dizziness, difficulty speaking and rapid heart rate (Stein and Stein, 2008). The excessive fear of embarrassment, humiliation and scrutiny results in social

isolation and functioning impairment to a significant degree (APA, 2013). The debilitating effect of SAD adversely affects a person's ability to work and school or rapport development. In addition to the impairment of education and occupation amongst individuals with SAD, a tremendous economic burden associated with SAD is incurred in society (Stuhldreher et al., 2014). The direct cost consists of healthcare costs due to outpatient treatment, hospitalisations, pharmaceuticals and formal and informal social services. The indirect cost includes reduced productivity due to sick leave and disability pension.

The common treatments for SAD involve medication, cognitive behavioural therapy (CBT), exposure therapy and group therapy (Acarturk et al., 2008; Blanco et al., 2003; Caletti et al., 2022). However, these therapies are limited by time and space and require the high commitment of patients in the whole process. Patients in cognitive–behavioural group therapy (CBGT) can obtain mutual support but managing a group of people is difficult for a therapist (Bouchard et al., 2017). The loss of privacy issue also emerges in CBGT. To overcome the drawbacks of these therapies, an alternative medium, virtual reality (VR), has recently been utilised in SAD therapy. VR is a computer-generated, lifelike virtual world that can immerse patients in a 3D environment and elicit their attention. Virtual reality therapy (VRT) in clinical settings makes patients encounter difficult situations and repeatedly experience such situations until they understand and learn from the scenarios (Deng et al., 2019). Potential barriers for patients who may have difficulty imagining or visualising can be eliminated by VR. VR allows for economical approaches and the potential to construct exposures that might be difficult to implement since in vivo exposures can be unfeasible to implement (e.g., terrorist attacks) and costly (e.g., natural disasters and plane crashes) (Maples-Keller et al., 2017). No input from a therapist may be required for some disorders. As such, VR can help improve access to the most effective psychotherapy (Freeman et al., 2017). Some studies have shown that the satisfaction level for VRT are higher than that for

traditional therapy (Brown and Foronda, 2020; Maples-Keller et al., 2017). In a study of Sarkar et al. (2022), older adults with knee osteoarthritis reported high levels of enjoyment of life, reduced mood, and pain intensity when mediated using VR.

Emotional processing theory (Foa and Kozak, 1986) can explain the effectiveness of VRT in changing the emotions and behaviours of an individual. This theory presumes that fear occurs when neural networks generate information about a stimulus, the meaning of the stimulus and the corresponding behaviour (Owens and Beidel, 2014). VRT therapy converts these relationships by generating new and incompatible information. To fulfil the principle of this theory, three conditions must be met for developing an appropriate VRT (Owens and Beidel, 2014). Firstly, the virtual environment should be made generally applicable to real-world situations. Secondly, the patients should immerse themselves in the virtual world. Thirdly, the virtual environment should induce physiological arousal, suggesting that core elements of fear are being addressed. A study has demonstrated the capability of VRT to provoke significant levels of subjective distress and immersion compared with in vivo exposure therapy (iVET) (Powers et al., 2013).

Effectiveness studies involving the comparison of VR-based therapy and traditional therapy (i.e. CBT) or waitlist control group have been conducted to assess the four mental health conditions associated with SAD: fear and avoidance of social interactions and performance situations (FASIP), fear of negative evaluation (FNE), anxiety (including social interaction anxiety, anxiousness, worry and social appearance anxiety) and depression. Mixed results have been identified for the effectiveness of VRT. For the first condition, extensive research has shown that VRT significantly reduces FASIP compared with traditional therapy (Bouchard et al., 2017; Denizci Nazligul et al., 2018; North et al., 1998; Robillard et al., 2010; Rubin et al., 2022; Safir et al., 2011) or waitlist control group (Harris et al., 2002). On the contrary, some studies have reported that the effect of VRT is similar to or not superior to

that of traditional therapy (Anderson et al., 2013; Anderson et al., 2017; Safir et al., 2011; Wallach et al., 2009; Wallach et al., 2011). For the second condition, several studies have demonstrated the significant effects of VRT on reducing FNE compared with traditional therapy (Bouchard et al., 2017; Denizci Nazligul et al., 2018; Robillard et al., 2010). However, some trials have demonstrated that VRT has no therapeutic efficacy in alleviating FNE (Anderson et al., 2017; Safir et al., 2011; Wallach et al., 2009; Wallach et al., 2011). For the third condition, a few studies have argued that VRT effectively reduced the symptoms of anxiety compared with traditional therapy (Bouchard et al., 2017; Klinger et al., 2005; North et al., 1998; Robillard et al., 2010; Rubin et al., 2022) or waitlist control group (Harris et al., 2002; Zanial et al., 2021), but a few trials found no difference between the effect of VRT and CBT on anxiety (Denizci Nazligul et al., 2018; Klinger et al., 2005). For the fourth condition, several studies have revealed a significant effect of VRT on depression compared with traditional therapy (Bouchard et al., 2017; Klinger et al., 2005; Robillard et al., 2010; Roy et al., 2003). In addition, the study of Kampmann et al. (2016) found that the therapeutic efficacy of iVET, a form of CBT, in improving FASIP, FNE, anxiety and depression is better than that of VRT. These inconsistent findings imply the necessity of conducting a meta-analysis to evaluate the effects of VRT on the mental health conditions associated with SAD. Although Horigome et al. (2020) conducted a meta-analysis of the effect of VRT on SAD, they examined only one outcome of SAD, focused on changes in the long-term period and included studies with different types of designs (e.g. cross-sectional, nonrandomised and observational). A growing number of randomised controlled trials (RCTs) have investigated the therapeutic efficacy of VRT for SAD in the recent decade; however, a meta-analysis of RCTs, which can provide convincing and robust evidence, is still lacking.

To our knowledge, no previous meta-analyses have categorised symptoms related to SAD into four mental health conditions, namely, FASIP, FNE, anxiety and depression to evaluate

the effects of VRT on SAD. FASIP and FNE are the major symptoms of SAD (Levinson et al., 2013), and more than 70% of patients with SAD are diagnosed with anxiety-depression comorbidity (Wong et al., 2014). Considering the inseparable relationships between SAD and these four mental health conditions, we aimed to conduct a systematic review and meta-analysis of RCTs that adopted VRT in clinical settings as an intervention for treating patients with SAD and compared the effects of VRT with CBT or waitlist control groups. In addition, the participant and study characteristics were identified to evaluate their moderating effects on the association between VRT and the four conditions. Four methodological features, namely, the headset used in the trials, the contents or tasks of VRT, the duration of intervention and acceptance of the use of VR, were comprehensively analysed to further ameliorate VRT in treating SAD. The results can identify whether VRT for treating SAD is presently ready for large-scale applications. Likewise, this article identifies features that are attributable to the effectiveness of VRT and that need to be refined to accommodate future VRT development in treating SAD. The findings can provide useful insights for professionals to optimise current therapies for patients and highlight new areas for investigation.

2. Method

2.1. Search strategy

This review focused on RCTs that used VRT in improving the symptoms of SAD and followed the PRISMA guidelines (Moher et al., 2009). This systematic review and meta-analysis was registered in the PROSPERO International Prospective Register of Systematic Reviews (REGISTRATION NUMBER: CRD42022346378). Medline, PubMed, Science Direct, Web of Science, CINAHL, PsychINFO and Scopus were searched up to May 2022. The search MeSH and keyword terms were as follows.

Population: MeSH terms: “social phobia” or “social phobias” or “social anxiety disorder” or “social anxiety” or “anxiety disorder”; keyword terms: “patients with social anxiety disorder” or “patients diagnosed with social anxiety disorder”

Intervention: MeSH terms:" virtual reality immersion therapy” or “virtual reality therapy” or ““virtual reality therapies”; keyword terms: “virtual reality” or “VR” or “virtual environment” or “virtual reality exposure therapy” or “virtual reality intervention” or “virtual reality cognitive-behaviour therapy” or “virtual reality exposure intervention”

Outcomes: “symptoms” or “fear” or “fear of negative evaluation” or “fear of public speaking” or “fear and avoidance of a range of social interactions and performance situations” or “anxiety” or “social interaction anxiety” or “anxiousness” or “depression”

2.2. Eligibility criteria

The included studies must follow the PICO framework, as follows: (a) population: patients who were diagnosed with SAD; (b) intervention: VR-related therapy or intervention for the purpose of improving the symptoms of SAD; (c) comparison: VR therapy versus waitlist control group, CBT or placebo-effect intervention; and (d) outcomes: FASIP, FNE, anxiety and depression. The other inclusion criteria were as follows: (a) the research design should be RCT and (b) the studies must be original and published in English. The exclusion criteria were (a) insufficient data for computing the effect sizes; (b) comorbidity with other mental health problems, such as affective disorder or autism spectrum disorder; and (c) case studies, cross-sectional studies, reviews, qualitative studies, observational studies, study protocols or conference abstracts without full text. Primary screening of the titles and abstracts was performed by two reviewers (K.P.W. and C.Y.Y.L.) to identify the relevant studies in the literature search. The references in the extracted studies were manually examined for eligibility. All corresponding authors for the unpublished data were contacted but none responded. After discarding the duplicates, the articles were independently scanned

by three reviewers (K.P.W., C.Y.Y.L and J.Q.), and full-text screening was conducted on the retrieved studies.

2.3. Data extraction and quality assessment

The following information of the reviewed studies was compiled when available: first author, year of publication, country, diagnosis, sample size, mean age, age range, the proportion of females in sample size, setting, trial arms, outcome measurements, software, length and duration of the therapy, length of follow-up, headset, attrition, acceptance and results. The mean and standard deviation (SD) were extracted for the effect size computation.

The risk of bias in the included studies was assessed using the Cochrane Collaboration tool (Higgins et al., 2011). The tool consists of seven criteria for judging the levels (i.e. high, low and unclear) of the risk of bias.

2.4. Data analysis

All data analyses were performed using the Comprehensive Meta-Analysis (CMA) Software Version 3.0. Hedge's g (*g*) with a 95% confidence interval (CI) was adopted in the analysis. The mean values, SDs and sample sizes in each trial in the post-test were extracted to compute the effect sizes. The random effect model was selected due to methodological diversity and study variability. Q and I^2 values were presented to measure the degree of heterogeneity. The Q value is closely related to the number of involved studies. Given its sensitivity to the number of studies, the Q value was not used as a reference for heterogeneity estimation in this study. Meanwhile, the greater the I^2 value is, the higher level of heterogeneity is. The I^2 value was zero in this study, indicating the absence of heterogeneity. In addition, trim and fill analysis and funnel plots using Egger's test were used to evaluate the publication bias. Subgroup analyses of different moderators were also conducted whenever available.

3. Results

3.1. Results of the literature search

Out of the 682 records extracted from the database search, 382 full-text articles were retrieved and screened after eliminating the duplicated articles. A total of 326 articles were excluded because they did not meet the selection criteria, and full-text screening was conducted on the remaining 56 studies. Fifteen RCTs met the inclusion and exclusion criteria. The flow chart of the article selection process is shown in Appendix A.

3.2. Participants' characteristics

The total sample included 720 participants, amongst which 76.6% were females. The average mean age was 31.0, and the age range was 18–69. The samples were from the United States (*n* = 6), Israel (*n* = 3), France (*n* = 2), Canada (*n* = 1), Turkey (*n* = 1) and the Netherlands (*n* = 1). To evaluate the symptoms of SAD amongst the participants, most studies adopted the structured clinical interview for DSM-IV (*n* = 9), followed by Public Speaking Anxiety Scale (*n* = 3), Liebowitz Social Anxiety Scale (LSAS) (*n* = 1), Personal Report of Confidence as a Speaker (*n* = 1) and Social Phobia Diagnostic Questionnaire (SPDQ) (*n* = 1). The details of the participants' characteristics are given in Table 1.

Table 1. Characteristics of the included studies

First author	Year	Country	Diagnosis	N	Mean age (SD)	Age range	Female n(%)	Setting	Trial arms (n)	Outcome measurement
Anderson	2013	The United States	SAD: DSM–IV	97	39.03 (11.26)	19–69	60 (62)	University	VRET (30); EGT (39); WL (28)	Fear of public speaking: PRCS; Fear of negative evaluation: BFNE
Anderson	2017	The United States	SAD: DSM-IV	65	42 (13.17)	19–69	46 (71)	University	VRET (33); EGT (32)	Fear of public speaking: PRCS; Fear of negative evaluation: BFNE
Bouchard	2017	Canada	SAD: DSM-IV	59	34.5 (11.9)	18–65	43 (72.9)	Laboratory of university	In virtuo+CBT (17); In vivo (22); WL (20)	Fear and avoidance of a range of social interactions and performance situations: LSAS-SR; Behaviour Avoidance: BAT; Social Interaction Anxiety: SIAS; Fear of negative evaluation: BFNE; Depression: BDI-II
Denizci Nazligul	2018	Turkey	LSAS	14	21.36 (1.08)	20–24	10 (71.4)	University	VREI (7); CBT (7)	Fear and avoidance of a range of social interactions and performance situations: LSAS; Anxiousness: IAS; Social appearance anxiety: SAAS; Fear of negative evaluation: BFNE
Harris	2002	The United States	PRCS	14	No mention (university student)	No mention	No mention	No mention	VRT (8); WL (6)	Fear of public speaking: PRCS; Anxiety: STAI; Fear and avoidance of a range of social interactions and performance situations: LSAS-SR
Kampmann	2016	The Netherlands	SAD: DSM-IV	60	36.9 (11.49)	18–65	38 (63.33)	Laboratory of the University	VRET (20); iVET (20); WL (20)	Fear and avoidance of a range of social interactions and performance situations: LSAS-SR; Fear of negative evaluation: BFNE;

										Depression: DASS depression; Anxiety: DASS Anxiety; Stress: DASS stress
Klinger	2005	France	SAD: DSM-IV	36	31.3 (7.91)	18–65	19 (52.78)	University Hospital	VRT (18); CBT (18)	Fear and avoidance of a range of social interactions and performance situations: LSAS-SR; Anxiety: SCIA; Anxiety and Depression: HAD
North	1998	The United States	SAD: DSM-IV	16	No mention (graduates)	No mention	No mention	University	VRT (8); TVRS (8)	Fear of public speaking: SUD; Anxiety: SUD
Robillard	2010	no mention	SAD: DSM-IV	45	34.9	No mention	32 (71)	No mention	In virtuo+CBT (14); CBT (16); WL (15)	Fear and avoidance of a range of social interactions and performance situations: LSAS; Anxiety: SAS; Fear of negative evaluation: BFNE; Depression: BDI II; Anxiety: STAI-Y2
Roy	2003	France	SAD: DSM-IV	10	36.1	18–65	6 (60)	Hospital	VRT (4); CBGT (6)	Fear and avoidance of a range of social interactions and performance situations: LSAS-SR; Anxiety: HAD-A; Depression: HAD-D, BDI-13
Rubin	2022	The United States	SAD: DSM-IV	21	22.7	18–65	13 (61.9)	No mention	VRET (10); VRET + AGT (11).	Social interaction anxiety: PRPSA; Fear and avoidance of a range of social interactions and performance situations: LSAS-SR Speech anxiety thoughts: SATI; Distress: SUDs; Columbia suicidality severity: C-SSRS

Safir	2011	Israel	PSA	49	27	No mention	70 (79)	No mention	VRCBT (25); CBT (24)	Fear and avoidance of a range of social interactions and performance situations: LSAS-SR; Fear of negative evaluation: BFNE
Wallach	2009	Israel	PSA	112	27	No mention	68 (77.3)	No mention	VRCBT (28); CBT (30); WL (30)	Fear and avoidance of a range of social interactions and performance situations: LSAS-SR; Fear of negative evaluation: BFNE
Wallach	2011	Israel	PSA	78	26.8	No mention	55 (70.5)	No mention	VRET (10); CT: (10); CBT (28); WL (30)	Fear and avoidance of a range of social interactions and performance situations: LSAS-SR; Fear of negative evaluation: BFNE
Zainal	2021	The United States	SPDQ	44	23.30 (9.32)	18–53	34 (77.3)	No mention	VRET (26); WL (18)	Social interaction anxiety: SAD, SIAS; Job interview anxiety: MASI; Worry: PSWQ

Note. AGT=attention guidance training; BAT=Behaviour Avoidance Test; BDI-II=Beck Depression Inventory-II; BFNE=Fear of negative evaluation, Brief Form; C-SSRS=Columbia Suicidality Severity Rating Scale; CBGT=cognitive-behavioural group therapy; CBT=cognitive-behaviour therapy; CT=cognitive therapy; DASS=Depression Anxiety Stress Scale; DSM-IV=Diagnostic and Statistical Manual of Mental Disorders, 4th Edition; EGT=Exposure Group Therapy; HAD=Hospital Anxiety and Depression; IAS=Interaction Anxiousness Scale; iVET=in vivo exposure therapy; LSAS-SR=Liebowitz Social Anxiety Scale; MASI=Measure of Anxiety in Selection Interviews; PRCS=Personal Report of Confidence as a Speaker; PRPSA=Personal Report of Public Speaking Anxiety; PSA=Public speaking anxiety; PSWQ=Penn State Worry Questionnaire; SAA=Social Appearance Anxiety Scale; SAD composite=averaged standardised score of the social interaction anxiety scale and social phobia diagnostic questionnaire; SAS=Social Anxiety Scale; SATI=Speech Anxiety Thoughts Inventory; SCIA=Social Contexts Inducing Anxiety; SIAS=Social Interaction Anxiety Scale; SPDQ=Social Phobia Diagnostic Questionnaire;

STAI=State-Trait Anxiety Inventory–State; SUD=Subjective Units of Disturbance Scale; TVRS=trivial virtual reality scene; VRCBT=Virtual reality cognitive-behaviour therapy; VREI=virtual reality exposure intervention; VRET=virtual reality exposure treatment; VRT=virtual reality therapy; WL=waitlist

3.3. Study characteristics

Out of the 15 studies, nine used a two-arm trial design and six used a three-arm design. Seven trials were conducted in a university and two in a hospital. The intervention group in all studies was subjected to VR exposure therapy. The control groups in the selected trials included the waitlist control group, exposure group therapy, iVET, CBT, trivial VR scene and attention guidance training. The primary task of the therapy was giving a speech in front of a group of people in a virtual environment; other tasks included attending a job interview, talking with friends or relatives and having a blind date. Most of the studies required the participants to wear a headset during the therapy; however, three trials did not report the hardware used in the VR therapy. The intervention duration varied amongst the studies and ranged within 1–14 weeks. The length of the sessions was heterogeneous and ranged within 10–60 minutes. The follow-up periods also differed amongst the studies, ranging within 0–72 months. The details of the interventions in each study are shown in Table 2.

3.4. Risk of bias

Four studies reported a low risk of random sequence generation. A low risk of allocation concealment was recorded in four studies. Two studies had a low risk of blinding of participants and personnel. Although two trials had no blinding of participants, they reported unavoidable broken blinding; thus, the two studies were regarded as having a low risk of blinding of participants and personnel. Four studies indicated the blinding of outcome assessment, and four studies reported appropriate methods for managing missing data. All studies were judged to have a low risk of selective reporting because all of them reported the primary and secondary outcomes even though several studies did not have the available protocol. Unclear risk of other bias was judged for all studies because all studies did not provide sufficient information to assess the existence of the related risk of bias. the details of the risk of bias are demonstrated in Appendix B.

3.5. Acceptability of VRT

The acceptability of VRT was evaluated based on the attrition rate. The average attrition rate was 9.30%. More than half of the trials reported no attrition in the whole intervention period (n = 8), and the highest attrition rate amongst the included studies was 25%. These results indicate that the acceptability of VRT amongst patients with SAD was high.

Table 2. Characteristics of the VRT of the included studies

First author (Year)	Software	Length	Follow-up	Headset	Attrition	Acceptance of VR (%)	Results
Anderson (2013)	Virtual environments of conference room, classroom and auditorium	8 weekly 20 min sessions	12-month	No mention	VRET (5/30); EGT (14/39); WL (3/28)	83.3	VRET was effective for treating social fears and improvement is maintained for 1 year.
Anderson (2017)	Virtual environments of conference room, classroom and auditorium	8 weekly sessions	72-month	No mention	VRET (0/33); EGT (0/32)	100	VRET and EGT for SAD produced long-lasting benefits.
Bouchard (2017)	Virtual environment: speaking in front of an audience in a meeting room, having a job interview, talking with apartment with supposed relatives, acting under the scrutiny of strangers on a coffee shop patio and facing criticism or insistence	14 weekly 20-30-min sessions	6-month	eMagin z800 head-mounted display and an InterSense Inertia Cube motion tracker	In virtuo+CBT (2/17); iVET (4/22); WL (4/20)	88.2	VR was significantly more practical for therapists than iVET.
Denizci Nazligul (2018)	Virtual environment of auditorium with three different sizes.	5 weekly 10-15-min sessions	No follow-up	HMD Oculus RIFT	VREI (0/7); CBT (0/7)	100	VREI was an alternative solution to the traditional interventions to overcome public presentation anxiety.
Harris (2002)	Virtual environment of auditorium	4 weekly 12-15-min sessions	Immediate follow-up after session 4	HMD with head-tracker (Virtual–I/0)	VRT (0/8); WL (0/6)	100	VRT was effective in reducing public speaking anxiety in university students.

Kampmann (2016)	Delft Remote Virtual Reality Exposure Therapy system, virtual environment: giving a talk in front of an audience followed by questions from the audience, talking to a stranger, buying and returning clothes, attending a job interview, being interviewed by journalists, dining in a restaurant with a friend, and having a blind date	10 90-min sessions in 5 weeks	3-month	nVisor SX head mounted display, stereographic projection	VRET (5/20); iVET (3/20); WL (2/20)	75	iVET was superior to the VRET regarding decreases in social anxiety symptoms at post- and follow-up assessments, and avoidant personality disorder related beliefs at follow-up. VRET was significant in reducing perceived stress only.
Klinger (2005)	Discreet 3D Studio Max 4, Virtools, virtual environment: speaking in front of a public audience, establishing contacts and talking with neighbors and friends, having small talk while being under scrutiny of unknown people, protecting his/her interests, viewpoints, and to be respected	12 weekly 20-min sessions	No mention	Mouse, arrow keys of the keyboard or a Cyberpuck	VRT (0/18); CBT (0/18)	100	VRT was statistically and clinically significant improvement in SAD and depression.
North (1998)	VREAM™ Virtual Reality Development Software Package and Libraries (VRCreator™), virtual environment of a list of public speaking situations	5 weekly 10-20 min sessions	No mention	Pentium-based™ computer (100 MHZ), HMD and Head-Tracker (Virtual I/O™)	VRT (2/8); TVRS (0/8)	75	VRT group showed significant improvement in reducing the fear of the public speaking after five weeks of treatment.

Robillard (2010)	European VERSY Updated project	16 weekly sessions	No mention	HP xw4600 Intel® Core™2 Duo E6850 3.00GHz 2GB/500GB, eMagin z800 3D Visor	In virtuo+CBT (0/14); CBT (0/16); WL (0/15)	100	In virtuo+CBT, CBT and WL had significant reduction of anxiety.
Roy (2003)	3D Studio Max 4 (3DS Max) graphic tool from Discreet, Virtools Dev 2.0, virtual environment of meeting rooms, kitchens, living room, coffee shop and shop	12 weekly 20-min sessions	No mention	No mention	VRT (0/4); CBGT (0/6)	100	VRT significantly reduced the anxiety and depression.
Rubin (2022)	Virtual environment of conference room and auditorium	2 45-min sessions	One-week	Oculus Rift DKII virtual reality headset	VRET (1/10); VRET + AGT (1/11).	90	Attention could be modified within and during VRET and that modification of visual gaze avoidance might be casually linked to reductions in social anxiety disorder.
Safir (2011)	Virtual environment with gradual exposure to aversive stimuli	12 weekly 60-min sessions	12-month	Computer, HMD (VFX3D from Interactive Imaging Systems)	VRCBT (0/25); CBT (0/24)	100	VRCBT was an effective and brief treatment regimen, equal to CBT. CBT continued to improve from post treatment to follow-up on LSAS fear.
Wallach (2009)	Virtual environment with gradual exposure to aversive stimuli	12 weekly	No mention	Computer, HMD (VFX3D from	VRCBT (6/28); CBT	78.6	VRCBT and CBT were significantly more effective than WL in anxiety reduction on four of five anxiety measures, and on

		60-min sessions		Interactive Imaging Systems)	(15/30); WL (3/30)		subject's self-rating of anxiety during a behavioural task. VRCBT is an effective and brief treatment regimen, equal to CBT.
Wallach (2011)	Virtual environment with gradual exposure to aversive stimuli	12 weekly 60-min sessions	No mention	Computer, HMD (VFX3D from Interactive Imaging Systems)	VRET (0/10); CT: (0/10); CBT (0/28); WL (0/30)	100	CT was not superior to VRET on cognitive measures but was superior to VRET on one behavioural measure (LSAS fear). VRET was superior to CT on one behavioural measure (fear reduction on a behavioural task). No differences were found between either CT, VRET and CBT, and all were superior to WL.
Zainal (2021)	Virtual environment of informal dinner party or the formal job interview	10 50-60 min sessions in 5 weeks	6-month	Pico Goblin VR headset	VRET (3/26); WL (2/18)	88.5	VRET resulted in greater reductions in SAD symptom severity, job interview fear, and trait worry, with moderate-to-large effect sizes from pre-to-post treatment. VRE led to change in depression, whereas waitlist did not.

Note. AGT=attention guidance training; CBGT=cognitive-behavioural group therapy; CBT=cognitive-behaviour therapy; CT=cognitive therapy; EGT=exposure group therapy; HMD=head-mounted display; iVET=in vivo exposure therapy; LSAS=Liebowitz Social Anxiety Scale; SAD=social anxiety disorder; Min=minutes; TVRS=Trivial virtual reality scene; VRCBT=virtual reality cognitive-behaviour therapy; VREI=virtual reality exposure intervention; VRET=virtual reality exposure treatment; VR=virtual reality; VRT=virtual reality therapy; WL=waitlist

3.6 Association between VRT and SAD

3.6.1. FASIP

Thirteen studies with 20 records evaluated the use of VRT for FASIP, with two records comparing with the waitlist control group and 18 records comparing with CBT. The measurements included LSAS (n = 15), Personal Report of Confidence as a Speaker (PRCS) (n = 3), Behavior Avoidance Test (BAT) (n = 1) and Subjective Units of Disturbance Scale (SUD) (n = 1). No significant effect was observed for FASIP when VRT compared with the control group (g = 0.005, SE = 0.124, 95% CI: -0.238–0.248; p < 0.969, Q = 41.5, I^2 = 54.2%). A large effect size for VRT versus the waitlist control group on FASIP was found (g = -1.170, SE = 0.452, 95% CI: -2.056–0.283; p < 0.010) and no significant difference between the effect of VRT and CBT on FASIP was observed (see Figure 1).

The results of the meta-regressions demonstrated that sample size moderated the effect size (≤50: g = -0.384, 95% CI: -0.729–0.038, p < 0.029; >50: g = 0.263, 95% CI: 0.004–0.522, p < 0.047). A mean age of 30 or higher demonstrated a significant moderating effect on effect size (g = 0.287, 95% CI: 0.041–0.532, p < 0.022). Number of sessions = ≤5 (g = -0.945, 95% CI: -1.485–0.405, p < 0.001) and number of sessions = 6–10 (g = 0.606, 95% CI: -0.285–0.926, p < 0.001) moderated the effect size (Table 3).

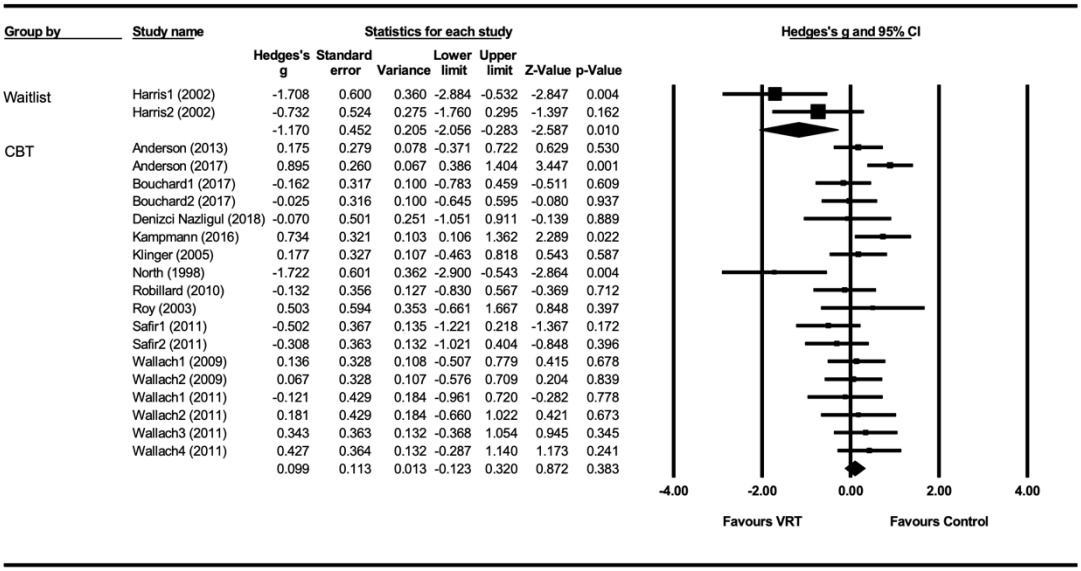

Figure 1. Forest plot showing Hedges' *g* and 95% confidence intervals (CIs) of studies on FASIP

Table 3. Meta-regressions for FASIP

Moderator	*n*	*g*	*SE*	95%CI	*z*	*p*
Year of publication						
1998-2010	8	-0.270	0.204	[-0.669; 0.129]	-1.328	0.184
2011-2022	12	0.160	0.149	[-0.133; 0.452]	1.069	0.285
Sample size						
≤50	9	-0.384	0.176	[-0.729; -0.038]	-2.177	0.029*
>50	11	0.263	0.132	[0.004; 0.522]	1.990	0.047*
Mean age						
≤30	9	0.026	0.134	[-0.238; 0.289]	0.190	0.849
>30	8	0.287	0.125	[0.041; 0.532]	2.289	0.022*
Females in sample size (%)						
Less	4	0.365	0.192	[-0.011; 0.741]	1.901	0.057
More	13	0.099	0.108	[-0.114; 0.311]	0.909	0.363
Virtual environment						
Giving a speech	5	-0.103	0.308	[-0.706; 0.501]	-0.333	0.739
More than giving a speech	5	0.223	0.294	[-0.352; 0.799]	0.760	0.447
Number of sessions						
≤5	4	-0.945	0.275	[-1.485; -0.405]	-3.431	0.001*
6-10	3	0.606	0.163	[0.285; 0.926]	3.706	0.000**
11-16	13	0.022	0.100	[-0.175; 0.218]	0.217	0.828
Session duration (min)						
≤30	9	-0.237	0.178	[-0.586; 0.112]	-1.330	0.184
>30	9	0.120	0.163	[-0.200; 0.439]	0.733	0.464
Acceptability						
Low	4	0.002	0.293	[-0.572; 0.577]	0.007	0.994
high	3	0.001	0.311	[-0.610; 0.611]	0.003	0.998
Very high	13	-0.007	0.168	[-0.337; 0.322]	-0.044	0.965

n: number of records, *g*: Hedge's *g*
* $p < 0.05$

** $p < 0.001$

3.6.2. FNE

Nine studies examined the difference between the effectiveness of VRT and CBT in FNE by using the Brief Fear of Negative Evaluation Scale (BFNE) (n = 10). No significant difference between the stimulus effect of VRT and CBT was found (g = 0.232, SE = 0.142, 95% CI: -0.046–0.51; $p < 0.102$, Q = 15.5, I^2 = 41.9%) (see Figure 2).

The results of the meta-regressions revealed that the sample size of more than 50 and session duration of more than 30 minutes had significant moderating effects on the effect size (sample size of > 50: g = 0.361, 95% CI: 0.071–0.651, $p < 0.015$; session duration of > 30 min: g = 0.444, 95% CI: 0.017–0.870, $p < 0.041$) (Table 2).

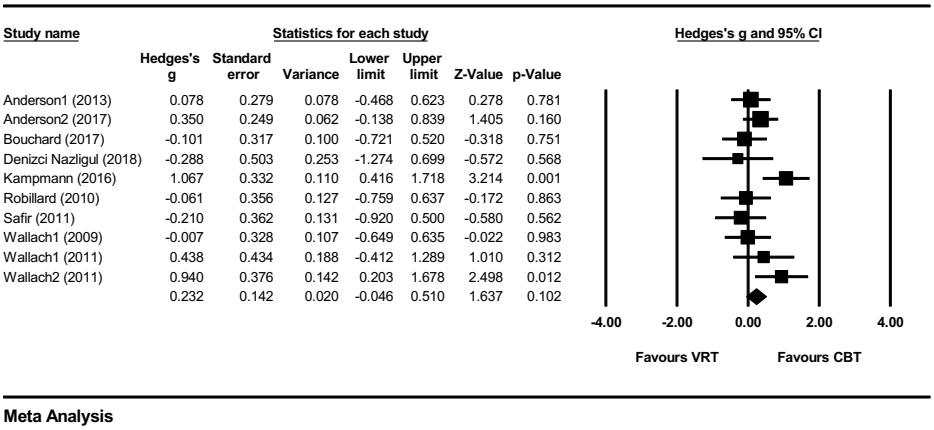

Figure 2. Forest plot showing Hedges' g and 95% confidence intervals (CIs) of studies on FNE

Table 4. Meta-regressions for FNE

Moderator	n	g	SE	95%CI	z	p
Year of publication						
1998-2010	2	-0.033	0.318	[-0.656; 0.590]	-0.014	0.917
2011-2022	8	0.299	0.160	[-0.014; 0.613]	1.870	0.061
Sample size						
≤50	3	-0.170	0.262	[-0.684; 0.345]	-0.647	0.518
>50	7	0.361	0.148	[0.071; 0.651]	2.442	0.015*
Mean age						
≤30	5	0.187	0.228	[-0.261; 0.634]	0.817	0.414
>30	5	0.266	0.199	[-0.124; 0.656]	1.336	0.182
Females in sample size (%)						
Less	2	0.527	0.296	[-0.054; 1.108]	1.777	0.076
More	8	0.144	0.162	[-0.173; 0.462]	0.890	0.374

Virtual environment						
Giving a speech	3	0.107	0.306	[-0.492; 0.706]	0.349	0.727
More than giving a speech	2	0.473	0.371	[-0.254; 1.200]	1.274	0.203
Number of sessions						
≤5	1	-0.288	0.581	[-1.427; 0.852]	-0.495	0.621
6-10	3	0.461	0.235	[0.000; 0.921]	1.959	0.050
11-16	6	0.140	0.189	[-0.231; 0.510]	0.739	0.460
Session duration (min)						
≤30	3	-0.066	0.275	[-0.606; 0.474]	-0.240	0.811
>30	5	0.444	0.218	[0.017; 0.870]	2.039	0.041*
Acceptability						
Low	2	0.526	0.321	[-0.103; 1.156]	1.638	0.101
high	2	-0.006	0.305	[-0.604; 0.591]	-0.020	0.984
Very high	6	0.220	0.198	[-0.167; 0.608]	1.115	0.265

n: number of records, *g*: Hedge's *g*
* $p < 0.05$

3.6.3. Anxiety

Nine studies (15 records) explored the efficacy of VRT in anxiety, with three records comparing with the waitlist control group and 12 records comparing with the CBT. The components associated with anxiety in the reviewed studies included social interaction anxiety, anxiousness, social appearance anxiety, job interview anxiety, worry and anxiety. The measures used were Social Interaction Anxiety Scale (SIAS) (*n* = 2), Interaction Anxiousness Scale (IAS) (*n* = 1), Social Appearance Anxiety Scale (SAAS) (*n* = 1), State-Trait Anxiety Inventory–State (STAI) (*n* = 2), Depression Anxiety Stress Scale (DASS) (*n* = 1), Social Contexts Inducing Anxiety (SCIA) (*n* = 1), Hospital Anxiety and Depression (HAD) (*n* = 2), SUD (*n* = 1), Social Anxiety Scale (SAS) (*n* = 1), SAD composite (*n* = 1), Measure of Anxiety in Selection Interviews (MASI) (*n* = 1) and Penn State Worry Questionnaire (PSWQ) (*n* = 1). No significant difference on anxiety was found when VRT compared with the control group (g = -0.264, SE = 0.168, 95% CI: -0.594–0.065; $p < 0.116$, Q = 41.115, I^2 = 65.949%). No difference between the effect of VRT in comparison with waitlist control group and CBT on anxiety was observed (see Figure 3).

The meta-regression for anxiety showed that year of publication, namely, 1998–2010, moderated the effect size (g = -0.526, 95% CI: -1.044; -0.008, $p < 0.05$). The sample size of

50 or smaller also moderated the effect size (g = -0.36, 95% CI: -0.709; -0.011, $p < 0.05$) (Table 5).

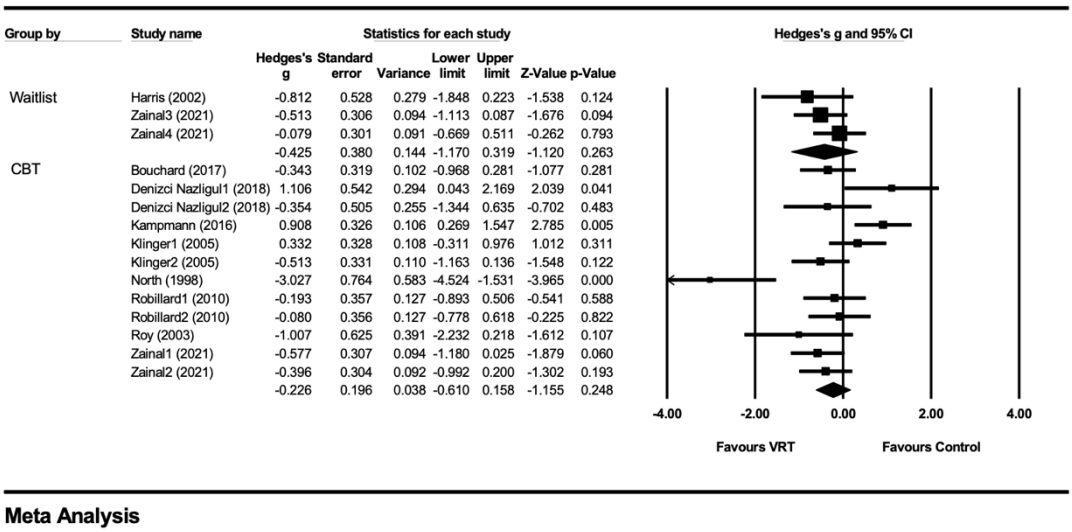

Figure 3. Forest plot showing Hedges' *g* and 95% confidence intervals (CIs) of studies on anxiety

Table 5. Meta-regressions for anxiety

Moderator	*n*	*g*	*SE*	95%CI	*z*	*p*
Year of publication						
1998-2010	7	-0.526	0.264	[-1.044; -0.008]	-1.990	0.047*
2011-2022	8	-0.075	0.228	[-0.521; 0.371]	-0.330	0.742
Sample size						
≤50	13	-0.360	0.178	[-0.709; -0.011]	-2.020	0.043*
>50	2	0.278	0.417	[-0.539; 1.096]	0.667	0.505
Mean age						
≤30	6	-0.228	0.216	[-0.651; 0.195]	-1.055	0.291
>30	7	-0.054	0.200	[-0.447; 0.338]	-0.272	0.785
Females in sample size (%)						
≤70	4	0.081	0.261	[-0.430; 0.592]	0.312	0.755
>70	9	-0.225	0.168	[-0.554; 0.104]	-1.341	0.180
Virtual environment						
Giving a speech	3	-0.037	0.403	[-0.827; 0.753]	-0.092	0.927
More than giving a speech	9	-0.204	0.191	[-0.579; 0.170]	-1.070	0.284
Number of sessions						
≤5	4	-0.573	0.402	[-1.361; 0.216]	-1.424	0.154
6-10	5	-0.138	0.287	[-0.700; 0.425]	-0.480	0.632
11-16	6	-0.252	0.277	[-0.795; 0.291]	-0.910	0.363
Session duration (min)						
≤30	8	-0.434	0.276	[-0.974; 0.107]	-1.573	0.116
>30	5	-0.137	0.307	[-0.739; 0.465]	-0.445	0.656
Acceptability						
Low	2	-0.316	0.520	[-1.334; 0.703]	-0.607	0.544
high	5	-0.381	0.276	[-0.922; 0.159]	-1.383	0.167
Very high	8	-0.167	0.244	[-0.646; 0.312]	-0.682	0.495

n: number of records, *g*: Hedge's *g*
* $p < 0.05$

3.6.4. Depression

Five studies with six records assessed the effect of between VRT and CBT on depression. The effects on depression were measured with Beck Depression Inventory-II (BDI-II) ($n = 3$), DASS ($n = 1$) and Hospital Anxiety and Depression (HAD) ($n = 2$). The Hedge's *g* for depression was 0.461 (95% CI: -0.49–0.582; $p < 0.866$, $Q = 14.4$, $I^2 = 65.3\%$), which was not at a statistically significant level (Figure 4).

The meta-regressions for depression revealed that '6–10 sessions', 'more than 30 min per session' and low acceptability seemed to moderate the effect size (Table 4). However, these categories of moderators only contained one record, so the statistical significance could not be established.

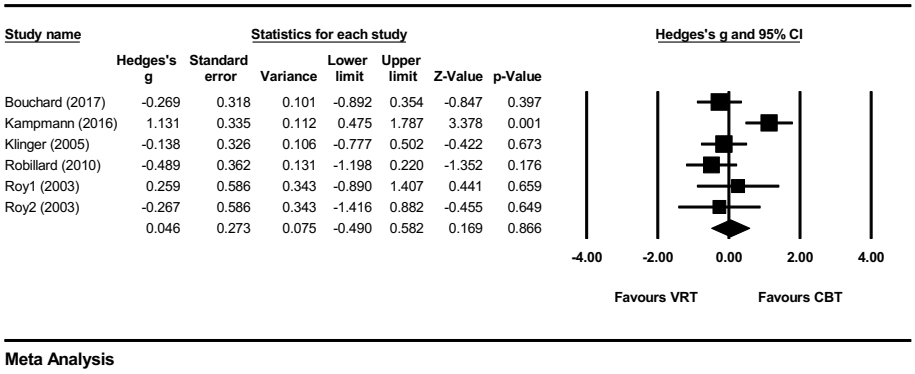

Figure 4. Forest plot showing Hedges' *g* and 95% confidence intervals (CIs) of studies on depression

Table 6. Meta-regressions for depression

Moderator	*n*	*g*	*SE*	95%CI	*z*	*p*
Year of publication						
1998-2010	4	-0.191	0.340	[-0.858; 0.476]	-0.560	0.575
2011-2022	2	0.420	0.427	[-0.417; 1.257]	0.984	0.325
Sample size						
≤50	4	-0.191	0.340	[-0.858; 0.476]	-0.560	0.575
>50	2	0.420	0.427	[-0.417; 1.257]	0.984	0.325
Mean age						
≤30	-	-	-	-	-	-
>30	6	0.046	0.273	[-0.490; 0.582]	0.169	0.866
Females in sample size (%)						
≤70	4	0.311	0.314	[-0.303; 0.926]	0.993	0.321

>70	2	-0.374	0.397	[-1.151; 0.403]	-0.943	0.346
Virtual environment						
Giving a speech	-	-	-	-	-	-
More than giving a speech	5	0.164	0.307	[-0.438; 0.767]	0.535	0.593
Number of sessions						
≤5	-	-	-	-	-	-
6-10	1	1.131	0.335	[0.475; 1.787]	3.378	0.001
11-16	5	-0.236	0.175	[-0.578; 0.107]	-1.349	0.177
Session duration (min)						
≤30	4	-0.158	0.200	[-0.550; 0.233]	-0.794	0.427
>30	1	1.131	0.335	[0.475; 1.787]	3.378	0.001
Acceptability						
Low	1	1.131	0.335	[0.475; 1.787]	3.378	0.001
high	-	-	-	-	-	-
Very high	5	-0.236	0.175	[-0.578; 0.107]	-1.349	0.177

n: number of records, *g*: Hedge's *g*

3.7. Publication bias

Publication bias was examined through trim and fill analysis (Duval and Tweedie, 2000) and Egger's test (Egger et al., 1997). An asymmetry shape in the funnel plots implies the existence of publication bias. The funnel plots for publication bias assessment are demonstrated in Appendix C. Five new data points were added to the funnel plots of FASIP (Appendix C1). One new data point was added to the funnel plots of FNE (Appendix C2). Two new data points were added to the funnel plots of anxiety (Appendix C3), and one new data point was added to the funnel plots of depression (Appendix C4). The results of Egger's test demonstrated that FASIP (z = -3.789, 95% CI: -6.064–1.515; $p < 0.005$) was statistically significant, suggesting the existence of publication bias. FNE, anxiety and depression were not statistically significant, indicating the absence of publication bias.

4. Discussion

To the best of our knowledge, this study is the first to evaluate the effectiveness of VRT in comparison with CBT and waitlist control group in treating the four psychological symptoms related to SAD, namely, FASIP, FNE, anxiety and depression amongst patients with SAD by meta-analytic review of RCTs. Contrary to expectations, this study located no significant differences between the effects of VRT and CBT on FASIP, FNE, anxiety and depression. These findings differ from the result of Horigome et al. (2020) that VRT had a

significant effect in treating SAD symptoms compared with in vivo exposure. There are several possible explanations for such a difference. The review of Horigome et al. (2020) did not restrict the types of study design, evaluated the outcomes measured by LSAS only and focused on the study characteristics in the moderator analysis. Our review included RCTs only, the outcomes related to SAD were categorised into four conditions (i.e. FASIP, FNE, anxiety and depression) and the coverage of the moderators included both sample characteristics and the arrangement and design of VRT. A meta-analysis of RCTs can provide a higher quality of evidence than a meta-analysis of empirical studies with different study designs (Moher and Olkin, 1995), which implied that our study may provide more reliable evidence than the study of Horigome et al. (2020). Further, evaluation of the effects of VRT on the four mental health conditions of patients with SAD could provide an in-depth understanding of the therapeutic efficacy of VRT for patients with SAD from different psychological aspects. By applying emotional processing theory, the meta-analytic results of no significant differences between the two groups may indicate the inapplicability of VRT to the real world, the low immersion amongst the participants during VRT and the incapability of inducing physiological arousal. Our study found that VRT outperformed the waitlist control group on FASIP. This finding is in line with the study of Carl et al. (2019) that VRT was effective in treating performance anxiety compared to the waitlist control group. Our narrative review indicated that VRT was partially effective in alleviating FASIP and FNE and reducing anxiety and depression. These findings suggest that the current level of immersion and interactivity of VRT still has room for improvement to enhance the therapeutic effect of VRT for SAD.

4.1. Quality of evidence

Amongst the 15 RCTs, only two (i.e. Anderson et al., 2013; Zainal et al., 2021) were rated as having moderate quality, and the remaining RCTs were of poor quality. Most of the

RCTs did not clearly report the random sequence generation and allocation concealment. Although broken blinding of participants was unavoidable during exposure therapy, many studies did not address how the participants were managed in the process. Most studies did not indicate the blinding of assessors. Only four RCTs stated the way they handled the attrition rate. All studies provided inadequate evidence on identifying the problems that may introduce bias. Not all studies provided an available protocol; however, all studies reported the primary and secondary outcomes. Therefore, additional high-quality and rigorously designed RCTs should be conducted to justify the effectiveness of VRT in SAD-related symptoms.

4.2. Moderator analysis

This study examined how the effect sizes of VRT differed across subgroups of studies on FASIP, FNE, anxiety and depression. Sample size, mean age, number of sessions, session duration and year of publication were found to be the moderators in the association amongst VRT, FASIP, FNE and anxiety. For FASIP, significant effects of VRT were observed amongst small sample sizes and small number of sessions. These may indicate that RCTs with few participants may benefit from VRT, and participants may exhibit considerable improvement with a low frequency of therapy. For FNE, the comparison group (i.e. CBT) had a small and significant effect amongst large sample sizes. This finding suggests that having many participants in a trial may favour the comparison group rather than VRT. For anxiety, a medium and significant effect of VRT was found amongst small sample sizes. This finding implies that having a few participants involved in the therapy may result in a considerable improvement. Surprisingly, a medium and significant effect of VRT was observed amongst the early publications. Recent publications should involve better-designed VRT compared with early publications due to the development or maturity of VR technology in recent years (Cipresso et al., 2018).

4.3. Methodological insights

The software platforms used for the VRT were varied, and the contents of the virtual environment differed across the RCTs. Five trials required the participants to give a speech in a conference room, classroom or auditorium, and five trials required the participants to conduct a series of social activities and public speaking situations, such as having a job interview; having small talks with different people in living rooms, coffee shops and stores; and having a blind date. However, five studies did not clearly elucidate the contents of their virtual environment. Therefore, additional public speaking situations need to be designed in virtual environments in lieu of giving a speech in front of many people. Patients with SAD desire acceptance from others and wish to leave a positive impression in their interactions (Beard et al., 2010). Meanwhile, they have a deep belief that their behaviours may attract criticism and negative attention, resulting in the avoidance of all such interactions (Hofmaan, 2007). VRT allows for more personalised therapy and controllable environments that might enhance the treatment effects (Bell et al., 2020; Sarkar et al., 2022). Effective therapy for patients with PTSD requires exposure to highly specific experiential cues, and customization of the virtual environment is a vital feature (Rizzo et al., 2010). VR can be adopted to evaluate how symptoms vary with different responses and triggers, such as sight, hearing, and smell. The determinants of events can be precisely identified through different settings in the virtual environment (Freeman et al., 2014; Veling et al., 2016). Given the advantages of VR, more lifelike interaction situations than the ones at present should be designed in virtual environments for patients to conquer social anxiety. Amongst the 15 RCTs, four did not specify the hardware (i.e. headset) adopted in the trials. The commonly used headsets were Oculus Rift, eMagin z800 3D Visor and VFX3D VR helmet. The absence of standardisation of the software and hardware used in the RCTs might possibly affect the significance of the effects of VRT on SAD.

Notably, having a few number of VRT sessions seemed to ameliorate the mental health symptoms associated with SAD. A possible explanation for this result might be that the participants could have been familiar with the virtual situation, so the effectiveness weakened in the subsequent sessions. Given that SAD is a mental health problem, the use of continuously repeated virtual environments might reduce the therapeutic efficiency. Numerous studies have shown that VRT is repetitive, allowing the therapist or practitioner to repeat a scene as many times as needed to help a patient (Garcia-Palacios et al., 2002; Giachero et al., 2020). These arguments have not been verified yet, so further research should be conducted to probe its validity. Furthermore, one unanticipated finding was that having a small sample size during the intervention was favourable for VRT. A possible explication for this might be that a large sample size may amplify the difference in detection and emphasise the statistical differences that are not clinically relevant.

The dropout rates of VRT were low across the studies. Eight RCTs reported full attendance for VRT. The attrition of VRT is high possibly due to the incapability for immersion in the VR environment, cybersickness or discomfort in the interaction (Chang et al., 2020; Weech et al., 2020). There is still room for investigating why VRT has low attrition.

4.4. Publication bias

The impact of publication bias on the pooling results is a major methodological concern in this study. A number of electronic databases, including Medline, PubMed, Science Direct, Web of Science, CINAHL, PsychINFO and Scopus, were used to retrieve the published articles. Google Scholar was adopted to explore grey literature, such as dissertations and theses, conference papers and government documents. Trim and fill analysis and Egger's test were applied to test the existence of publication bias. Our study found that the results of FNE,

anxiety and depression were not biased; however, the result of FASIP was affected by publication bias.

4.5. Limitations and future research directions

Several limitations of this review need to be acknowledged. Firstly, no significant difference was found between VRT and control groups in terms of the four mental health conditions of SAD, namely, FASIP, FNE, anxiety and depression. This absence made the summarised effect sizes highly complex. Additional RCTs should be conducted on the effects of VRT on SAD in the future. Secondly, each RCT should entail a comprehensive mental health condition assessment rather than focusing on the assessment of one symptom because relatively few studies examined the effect of VRT on depression (n = 5). Thirdly, although numerous potential moderators were identified, most were found to be of no significance. A multi-armed factorial design should be adopted to explore the methods for enhancing the efficacy of VRT. Ultimately, each study varied considerably in the hardware and software used. Changes in technology may affect the impact of VRT on patients.

The hardware and software platforms used for VRT development should be explicitly stated, which may assist in identifying the most comfortable and suitable equipment for patients when conducting VRT. Three studies did not mention the models of used headsets, and one of the included studies (i.e. Klinger et al., 2005) used a large screen to display the virtual environment in lieu of HMD. To maintain the consistency of the whole study, HMD with a similar quality level should be used in all VRTs. Furthermore, most trials were conducted at universities, laboratories and hospitals. Only a few non-randomised controlled trials examined the effectiveness of the mobile- or home-based VRT in patients with SAD (e.g. Kim et al., 2017; Nason et al., 2023) and thus more home- or mobile-based interventions are recommended for future RCT due to its convenience and cost-effectiveness. A home- or mobile-based intervention may increase the frequency of conducting VRT so that the efficacy

of the high frequency of conducting VRT in improving the symptoms associated with SAD can be assessed. For instance, Jeong et al. (2021) suggested that at least nine sessions of VRT may be effective in improving FNE. In addition, the duration of each session should be standardised to appraise the effectuality of VRT efficaciously and unbiasedly in treating patients with SAD. A 20-minute VRT session is recommended for future RCTs, as prolonged use of HMDs may lead to VR sickness (Taubert et al., 2019). Munafo et al. (2017) showed that the incidence of motion sickness was higher if using VR for more than 30 minutes. Therefore, each VRT should be designed within a safe duration. Lastly, this meta-analysis only covered English-language publications, so those published in non-English languages may have been missed.

5. Conclusion

The present meta-analysis found no statistically significant difference between the effect of VRT and CBT on the four mental health conditions related to SAD, namely, FASIP, FNE, anxiety and depression, and a large effect size for VRT versus the waitlist control group on FASIP. However, the narrative review showed that VRT partially alleviated the four conditions in comparison with CBT. Sample size, mean age, number of sessions, session duration and year of publication year were the potential moderators in the relationship between VRT, FASIP, FNE and anxiety. The results suggest the following: standardisation of VRT contents and headsets used, the concern regarding few therapy sessions having a significant effect on reducing the symptoms, adopting an appropriate VRT session length, and the possible influencing effects of participants' age on the dropout rate of VRT. Practitioners and researchers involved in designing VRT for patients with SAD should not only emphasise the software and hardware used for VRT but also be aware of the contents of the therapy and the number and duration of each session, which might possibly affect the outcomes.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors declare no conflict of interest.

References

Acarturk, C., Cuijpers, P., van Straten, A., de Graaf, R., 2008. Psychological treatment of social anxiety disorder: a meta-analysis. Psychol. Med. 39(02), 241-254.

American Psychiatric Association., 2013. Diagnostic and statistical manual of mental disorders. 5th ed. Washington, DC.

Anderson, P. L., Edwards, S. M., Goodnight, J. R., 2017. Virtual reality and exposure group therapy for social anxiety disorder: Results from a 4–6 year follow-up. Cogn. Behav. Ther. 41(2), 230–236.

Anderson, P. L., Price, M., Edwards, S. M., Obasaju, M. A., Schmertz, S. K., Zimand, E., Calamaras, M. R., 2013. Virtual reality exposure therapy for social anxiety disorder: A randomized controlled trial. J. Consult. Clin. Psychol. 81(5), 751.

Beard, C., Moitra, E., Weisberg, R. B., Keller, M. B., 2010. Characteristics and predictors of social phobia course in a longitudinal study of primary-care patients. Depression Anxiety, 27(9), 839–845.

Bell, I. H., Nicholas, J., Alvarez-Jimenez, M., Thompson, A., Valmaggia, L., 2020. Virtual reality as a clinical tool in mental health research and practice. Dialogues Clin. Neurosci. 22(2), 169–177.

Blanco, C., Schneier, F. R., Schmidt, A., Blanco-Jerez, C.-R., Marshall, R. D., Sánchez-Lacay, A., Liebowitz, M. R., 2003. Pharmacological treatment of social anxiety disorder: A meta-analysis. Depression Anxiety, 18(1), 29–40.

Bouchard, S., Dumoulin, S., Robillard, G., Guitard, T., Klinger, É., Forget, H., … Roucaut, F. X., 2017. Virtual reality compared with in vivo exposure in the treatment of social anxiety disorder: A three-arm randomised controlled trial. Br. J. Psychiatry 210(04), 276–283.

Brown, K., Foronda, C., 2020. Use of virtual reality to reduce anxiety and pain of adults undergoing outpatient procedures. Informatics, 7(3), 36.

Caletti, E., Massimo, C., Magliocca, S., Moltrasio, C., Brambilla, P., Delvecchio, G. 2022. The role of the acceptance and commitment therapy in the treatment of social anxiety: An updated scoping review. J. Affect. Disord. 310, 174–182. https://doi.org/10.1016/j.jad.2022.05.008

Carl, E., Stein, A. T., Levihn-Coon, A., Pogue, J. R., Rothbaum, B., Emmelkamp, P., Asmundson, G., Carlbring, P., Powers, M. B. 2019. Virtual reality exposure therapy for anxiety and related disorders: A meta-analysis of randomized controlled trials. J. Anxiety Disord. 61, 27–36. https://doi.org/10.1016/j.janxdis.2018.08.003

Chang, E., Kim, H. T., Yoo, B., 2020. Virtual reality sickness: A review of causes and measurements. Int. J. Hum.-Comput. Interact. 36(17), 1658–1682.

Cipresso, P., Giglioli, I. A. C., Raya, M. A., Riva, G., 2018. The past, present, and future of virtual and augmented reality research: A network and cluster analysis of the literature. Front. Psychol. 9. https://doi.org/10.3389/fpsyg.2018.02086

Deng, W., Hu, D., Xu, S., Liu, X., Zhao, J., Chen, Q., Liu, J., Zhang, Z., Jiang, W., Ma, L., Hong, X., Cheng, S., Liu, B., Li, X. 2019. The efficacy of virtual reality exposure therapy for PTSD symptoms: A systematic review and meta-analysis. J. Affect. Disord. 257, 698–709. https://doi.org/10.1016/j.jad.2019.07.086

Denizci Nazligul, M., Yilmaz, M., Gulec, U., Yilmaz, A. E., Isler, V., O'Connor, R. V., ... Clarke, P., 2019. An interactive 3d virtual environment to reduce the public speaking anxiety levels of novice software engineers. IET Software, 13(2), 152–158.

Duval, S., Tweedie, R., 2000. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 56(2), 455–463.

Foa, E. B., Kozak, M. J., 1986. Emotional processing of fear: exposure to corrective information. Psychol. Bull. 99(1), 20–35.

Freeman, D., Evans, N., Lister, R., Antley, A., Dunn, G., Slater, M., 2014. Height, social comparison, and paranoia: an immersive virtual reality experimental study. Psychiatry Res. 218(3), 348–352.

Freeman, D., Reeve, S., Robinson, A., Ehlers, A., Clark, D., Spanlang, B., Slater, M., 2017. Virtual reality in the assessment, understanding, and treatment of mental health disorders. Psychol. Med. 47(14), 2393–2400.

Garcia-Palacios, A., Hoffman, H., Carlin, A., Furness, T. A., Botella, C., 2002. Virtual reality in the treatment of spider phobia: a controlled study. Behav. Res. Ther. 40(9), 983–993.

Giachero, A., Calati, M., Pia, L., La Vista, L., Molo, M., Rugiero, C., Fornaro, C., Marangolo, P., 2020. Conversational therapy through semi-immersive virtual reality environments for language recovery and psychological well-being in post stroke aphasia. Behavioural neurology, Behav. Neurol. 2846046.

Gilbert, P. 2001., Evolution and social anxiety: The role of attraction, social competition, and social hierarchies. Psychiatr. Clin. North Am. 24(4), 723–751. https://doi.org/10.1016/S0193-953X(05)70260-4

Harris, S. R., Kemmerling, R. L., North, M. M., 2002. Brief virtual reality therapy for public speaking anxiety. Cyberpsychol. Behav. 5(6), 543–550.

Higgins, J. P., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., Savović, J., Schulz, K. F., Weeks, L., Sterne, J. A., Cochrane Bias Methods Group Cochrane Statistical Methods Group., 2011. The cochrane collaboration's tool for assessing risk of bias in randomised trials. BMJ 343, d5928.

Hofmann, S. G., 2007. Cognitive factors that maintain social anxiety disorder: a comprehensive model and its treatment implications. Cogn. Behav. Ther. 36(4), 193–209. doi:10.1080/16506070701421313

Horigome, T., Kurokawa, S., Sawada, K., Kudo, S., Shiga, K., Mimura, M., Kishimoto, T., 2020. Virtual reality exposure therapy for social anxiety disorder: a systematic review and meta-analysis. Psychol. Med. 50(15), 2487–2497.

Jeong, H. S., Lee, J. H., Kim, H. E., Kim, J-J., 2021. Appropriate number of treatment sessions in virtual reality-based individual cognitive behavioral therapy for social anxiety disorder. J. Clin. Med. 10(5), 915.

Kampmann, I. L., Emmelkamp, P. M., Morina, N., 2016b. Meta-analysis of technology-assisted interventions for social anxiety disorder. J. Anxiety Disord. 42, 71–84.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., Walters, E. E., 2005. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. Arch. Gen. Psychiatry 62(6), 593.

Kim, H. E., Hong, Y. J., Kim, M. K., Jung, Y. H., Kyeong, S., Kim, J. J., 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.

Klinger, E., Bouchard, S., Légeron, P., Roy, S., Lauer, F., Chemin, I., Nugues, P., 2005. Virtual reality therapy versus cognitive behavior therapy for social phobia: A preliminary controlled study. Cyberpsychol. Behav. 8(1), 76–88.

Levinson, C. A., Rodebaugh, T. L., White, E. K., Menatti, A. R., Weeks, J. W., Iacovino, J. M., Warren, C. S., 2013. Social appearance anxiety, perfectionism, and fear of negative evaluation: distinct or shared risk factors for social anxiety and eating disorders?. Appetite, 67, 125–133.

Maples-Keller, J. L., Bunnell, B. E., Kim, S.-J., Rothbaum, B. O., 2017. The Use of Virtual Reality Technology in the Treatment of Anxiety and Other Psychiatric Disorders. Harv. Rev. Psychiatry 25(3), 103–113.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., 2009. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. PLoS Med. 6(7), e1000097.

Moher, D, Olkin, I., 1995. Meta-analysis of randomized controlled trials: A concern for standards. J. Am. Med. Assoc. 274(24), 1962–1964. doi:10.1001/jama.1995.03530240072044

Munafo, J., Diedrick, M., Stoffregen, T. A., 2017. The virtual reality head-mounted display Oculus Rift induces motion sickness and is sexist in its effects. Exp. Brain Res. 235, 889–901.

Nason, E. E., Trahan, M., Cash, D., 2023. Evaluation of a mobile virtual reality intervention for social anxiety disorder: Ethical and methodological lessons learned. J. Technol. Behav. Sci. 8, 79–86. https://doi.org/10.1007/s41347-022-00292-0

North, M. M., North, S. M., Coble, J. R., 1998. Virtual reality therapy: An effective treatment for the fear of public speaking. Int. J. Virtual Real. 3(3), 1–6.

Owens, M. E., Beidel, D. C., 2014. Can virtual reality effectively elicit distress associated with social anxiety disorder? J. Psychopathol. Behav. Assess. 37(2), 296–305.

Powers, M. B., Briceno, N. F., Gresham, R., Jouriles, E. N., Emmelkamp, P. M., Smits, J. A., 2013. Do conversations with virtual avatars increase feelings of social anxiety? J. Anxiety Disord. 27(4), 398–403.

Rizzo, A. S., Difede, J., Rothbaum, B. O., Reger, G., Spitalnick, J., Cukor, J., McLay, R., 2010. Development and early evaluation of the Virtual Iraq/Afghanistan exposure therapy system for combat-related PTSD. Ann. N. Y. Acad. Sci. 1208, 114–125. https://doi.org/10.1111/j.1749-6632.2010.05755.x

Robillard, G., Bouchard, S., Fournier, T., Renaud, P., 2003. Anxiety and presence during VR immersion: A comparative study of the reactions of phobic and non-phobic participants in therapeutic virtual environments derived from computer games. Cyberpsychol. Behav. 6(5), 467–476.

Roy, S., Klinger, E., Légeron, P., Lauer, F., Chemin, I., Nugues, P., 2003. Definition of a VR-based protocol to treat social phobia. Cyberpsychol. Behav. 6(4), 411–420.

Rubin, M., Muller, K., Hayhoe, M. M., Telch, M. J., 2022. Attention guidance augmentation of virtual reality exposure therapy for social anxiety disorder: a pilot randomized controlled trial. Cogn. Behav. Ther. 6, 1-17.

Safir, M. P., Wallach, H. S., Bar-Zvi, M., 2012. Virtual reality cognitive- behavior therapy for public speaking anxiety: One-year follow-up. Behav. Modif. 36(2), 235–246.

Sarkar, T. D., Edwards, R. R., Baker, N., 2022. The feasibility and effectiveness of virtual reality meditation on reducing chronic pain for older adults with knee osteoarthritis. Pain Pract. 22(7), 631–641.

Stein, M. B., Stein, D. J., 2008. Social anxiety disorder. Lancet, 371, 1115–1125.

Stuhldreher, N., Leibing, E., Leichsenring, F., Beutel, M. E., Herpertz, S., Hoyer, J., … König, H.-H., 2014. The costs of social anxiety disorder: The role of symptom severity and comorbidities. J. Affect. Disord. 165, 87–94. doi:10.1016/j.jad.2014.04.039

Taubert, M., Webber, L., Hamilton, T., Carr, M., Harvey, M., 2019. Virtual reality videos used in undergraduate palliative and oncology medical teaching: results of a pilot study. BMJ Support. Palliat. Care. 9(3), 281–285.

Veling, W., Pot-Kolder, R., Counotte, J., van Os, J., & van der Gaag, M., 2016. Environmental social stress, paranoia and psychosis liability: A virtual reality Study. Schizophr. Bull. 42(6), 1363–1371.

Vilaplana-Pérez, A., Pérez-Vigil, A., Sidorchuk, A., Brander, G., Isomura, K., Hesselmark, E., … Fernández de la Cruz, L., 2020. Much more than just shyness: the impact of social anxiety disorder on educational performance across the lifespan. Psychol. Med. 51, 861-869.

Wallach, H. S., Safir, M. P., Bar-Zvi, M., 2009. Virtual reality cognitive behavior therapy for public speaking anxiety: A randomized clinical trial. Behav. Modif. 33(3), 314–338.

Wallach, H. S., Safir, M. P., Bar-Zvi, M., 2011. Virtual reality exposure versus cognitive restructuring for treatment of public speaking anxiety: A pilot study. Isr. J. Psychiatry Relat. Sci. 48(2), 91–97.

Wong, J., Morrison, A. S., Heimberg, R. G., Goldin, P. R., Gross, J. J., 2014. Implicit associations in social anxiety disorder: the effects of comorbid depression. J. Anxiety Disord. 28(6), 537–546.

Zainal, N. H., Chan, W. W., Saxena, A. P., Taylor, C. B., Newman, M. G., 2021. Pilot randomized trial of self-guided virtual reality exposure therapy for social anxiety disorder. Behav. Res. Ther. 147, 103984. https://doi.org/10.1016/j.brat.2021.103984