ELSEVIER

Journal of Anxiety Disorders 22 (2008) 561–569

Journal of Anxiety Disorders

Review

Virtual reality exposure therapy for anxiety disorders: A meta-analysis

Mark B. Powers *, Paul M.G. Emmelkamp

University of Amsterdam, The Netherlands

Received 1 March 2007; received in revised form 11 April 2007; accepted 20 April 2007

Abstract

There is now a substantial literature investigating virtual reality exposure therapy (VRET) as a viable treatment option for anxiety disorders. In this meta-analysis we provide effect size estimates for virtual reality treatment in comparison to in vivo exposure and control conditions (waitlist, attention control, etc.). A comprehensive search of the literature identified 13 studies ($n = 397$) that were included in the final analyses. Consistent with prediction the primary random effects analysis showed a large mean effect size for VRET compared to control conditions, Cohen's $d = 1.11$ (S.E. = 0.15, 95% CI: 0.82–1.39). This finding was consistent across secondary outcome categories as well (domain-specific, general subjective distress, cognition, behavior, and psychophysiology). Also as expected in vivo treatment was not significantly more effective than VRET. In fact, there was a small effect size favoring VRET over in vivo conditions, Cohen's $d = 0.35$ (S.E. = 0.15, 95% CI: 0.05–0.65). There was a trend for a dose–response relationship with more VRET sessions showing larger effects ($p = 0.06$). Outcome was not related to publication year or sample size. Implications are discussed.

Keywords: Virtual reality; Exposure therapy; Anxiety disorders; Meta-analysis

Contents

* Corresponding author at: Faculty of Social and Behavioral Sciences, Clinical Psychology, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands. Tel.: +31 20 525 6039; fax: +31 20 639 1369.

E-mail address: mpowers@fmg.uva.nl (M.B. Powers).

0887-6185/$ – see front matter
doi:10.1016/j.janxdis.2007.04.006

Exposure-based treatments for anxiety disorders show some of the largest effect sizes in the literature (Deacon & Abramowitz, 2004; Eddy, Dutra, Bradley, & Westen, 2004; Gould, Otto, & Pollack, 1995; Gould, Otto, Pollack, & Yap, 1997; Kobak, Greist, Jefferson, Katzelnick, & Henk, 1998). However, many patients are reticent to seek out exposure-based treatment. A recent development in the behavioral treatment of specific phobias is providing exposure through virtual reality (VR). Clients are not confronted with real anxiety provoking stimuli but with their virtual counterparts. Virtual reality integrates real-time computer graphics, body tracking devices, visual displays and other sensory input devices to immerse patients in a computer-generated virtual environment. One survey of students with a fear of spiders showed that almost 90% would prefer VR exposure over in vivo exposure therapy (Garcia-Palacios, Hoffman, See, Tsai, & Botella, 2001). This may represent a tendency toward avoidance of "real" feared stimuli. However, virtual reality exposure therapy (VRET) may encourage patients to get help who otherwise may not. Originally, most research on VRET as stand alone treatment was based on case studies, but more recently a number of randomized, controlled studies have been conducted, which strengthen the conclusions on the effectiveness of VRET (Emmelkamp, 2005; Krijn, Emmelkamp, Olafsson, & Biemond, 2004). However, to date no meta-analytic studies have investigated the combined effect of the large number of controlled studies that have accumulated.

Controlled studies to date show VRET may be an effective exposure delivery method for treating panic disorder (Botella et al., in press), social phobia (Harris, Kemmerling, & North, 2002; Klinger et al., 2005), PTSD (Difede et al., in press), fear of flying (Krijn et al., 2007; Maltby, Kirsch, Mayers, & Allen, 2002; Rothbaum et al., 2006; Rothbaum, Hodges, Smith, Lee, & Price, 2000), fear of spiders (Garcia-Palacios, Hoffman, Carlin, Furness, & Botella, 2002), and fear of heights (Emmelkamp et al., 2002; Krijn, Emmelkamp, Biemond et al., 2004; Rothbaum et al., 1995). Further, in these studies VRET was effective across multiple assessment domains including domain-specific subjective distress, general subjective distress, cognitive, behavioral, and psychophysiological measures. In comparing VRET with existing treatments studies suggest VRET may be equipotent to in vivo exposure with no significant differences between these two conditions (Rothbaum et al., 2006).

The current study employs a meta-analytic approach to test several hypotheses derived from the extant literature. Hypothesis 1: virtual reality exposure therapy would outperform control conditions on fear-specific measures. Hypothesis 2: results would be consistent across secondary outcome variables (general subjective distress, cognition, behavior, and psychophysiology). Hypothesis 3: there would be no significant difference between virtual reality exposure therapy and in vivo conditions. We conservatively set alpha at 0.25 or greater for this comparison. In addition, although not based on the current literature we conducted explorative analyses for the following research questions. Hypothesis 4: there would be a dose–response relationship for VRET. Hypothesis 5: sample size and publication year would not moderate the overall effect size.

1. Method

1.1. Study selection

We selected well-controlled trials with random or matched assignment of virtual reality exposure therapy for anxiety disorders using a comprehensive search strategy. We searched the following databases: PsycINFO 1840 to February, 2007, MEDLINE 1966 to February, 2007, and the Cochraine Central Register of Controlled Trials for the first quarter of 2007. The searches included the following terms: "virtual reality" alone and in combination with "exposure", "treatment" and "therapy". These words were searched as key words, title, abstract, and MeSH subject heading terms. Also, we examined citation maps and used the "cited by" search tools. These findings were cross-referenced with references from reviews. These initial

search strategies identified 95 potential articles. Next we limited the findings to human and English language studies. Studies meeting the following inclusion criteria were selected for the meta-analysis: (a) at least one virtual reality exposure therapy condition, (b) random assignment or matched conditions, and (c) either an active or inactive control group. Authors of selected studies were contacted directly when there were insufficient data provided in their articles to include in the meta-analysis. Thirteen studies with a total sample size of 397 participants met the final inclusion criteria and were included in the meta-analysis. The following studies were not included for a variety of reasons. The Emmelkamp, Bryunzeel, Drost, and van der Mast (2001) and Krijn, Emmelkamp, Olafsson, Schuemie, and van der Mast (in press) (both on acrophobia) studies were excluded, since these studies used a within and cross-over design, respectively, rather than a between group design (Emmelkamp et al., 2001; Krijn et al., in press). The Muhlberger, Wiedeman, and Pauli (2003) study on fear of flying and the Vincelli et al. (2003) study on panic disorder were not included since VRET was not investigated as stand alone treatment, but was combined with cognitive interventions, so that effects of VRET alone could not be established (Muhlberger et al., 2003; Vincelli et al., 2003). Finally, the Ressler et al. (2004) study on acrophobia was excluded since VR was combined with medication or placebo but a treatment control group was not included (Ressler et al., 2004). Given there was only one study comparing VRET with imaginal exposure the Wiederhold et al. (2002) was not included (Wiederhold et al., 2002). An additional study of VRET for public speaking (North, North, & Coble, 1998) was excluded because the data for meta-analysis were not available (personal communication, February, 2007).

1.2. Software

All analyses were completed with comprehensive meta-analysis (Borenstein & Rothstein, 1999). Comprehensive meta-analysis is a program funded by the National Institutes of Health SBIR program.

1.3. Procedure

Data on the following variables were collected: sample source (clinical/analogue), treatment dose (number of sessions), assignment (random/matched), number of participants per condition, disorder, and year of publication. Dependent variables were classified into categories including: domain-specific subjective distress (e.g. LSAS, PDSS, etc.), general subjective distress (e.g. BDI, SCL-90, etc.), cognitive (e.g. ATHQ, ATPS, etc.), behavioral (e.g. behavioral approach, BATs), and psychophysiological (e.g. HR, SCL, etc.) (see Table 1).

Treatments were categorized into the following treatment condition ''types'': virtual reality exposure therapy (VRET) or in vivo exposure. Control conditions were categorized into one of the following: waitlist (WL), relaxation, bibliotherapy, or attention placebo.

1.4. Effect size calculation

Between-group effect sizes for each study were computed using Hedge's g (Rosenthal, 1991). Studies with multiple outcomes were categorized as above and then combined within each domain. When the necessary data were available, all effect sizes were calculated directly using the following formula: $g = \bar{X}_T - \bar{X}_C / S_P$ where $\bar{X}_T$ is the mean of the treatment group, $\bar{X}_C$ the mean of the comparison group, and S_P is the pooled standard deviation. If these data were not provided, g

Table 1
Measure coding for analyses

Domain	Measure
Domain-specific subjective distress	AES, AQ, CAPS, CGI, DE, FAM-somatic, FAS-generalized, FFI, FFS, FOFR, FA, FSQ, FQ, LSAS, PDSS, PRCS, QAF, SUDs,
General subjective distress	ASI, BDI, BSI, STAI, HAD-A, HAD-D, SDS
Cognitive	ATHQ, ATPS, FAM-cognitive
Behavioral	BAT, behavioral approach, flights
Psychophysiological	HR, SCL

AES: anxiety expectancy scale; ASI: anxiety sensitivity index; ATHQ: attitude towards height questionnaire; ATPS: attitudes towards public speaking questionnaire; AQ: anxiety questionnaire and acrophobia questionnaire; BAT: behavioral approach test; BDI: beck depression inventory; BSI: brief symptom inventory; CAPS: clinician administered PTSD scale; CGI: clinical global impression; DE: danger expectancy; FA: flying avoidance; FAM: flight anxiety modality questionnaire; FAS: flight anxiety situations questionnaire; FFI: fear of flying inventory; FFS: fear of flying scale; FOFR: fear of flying rating; FSQ: fear of spiders questionnaire; FQ: fear questionnaire; HAD: hospital anxiety and depression; HR: heart rate; LSAS: Liebowitz social anxiety scale; PDSS: panic disorder severity scale; PRCS: person report of confidence as a speaker; QAF: questionnaire on attitudes toward flying; SCL: skin conductance level; SDS: Sheehan disability scale; SUDs: subjective units of distress.

was estimated using conversion equations for significance tests (e.g. t, F; see Rosenthal, 1991). All effect sizes were corrected for small sample sizes according to Hedges and Olkin (1985). Therefore, a smaller sample size reduces the estimated effect size helping control for different sample sizes across studies. These controlled effect sizes may then be conservatively interpreted with Cohen's (1988) convention of small (0.2), medium (0.5), and large (0.8) effects (Cohen, 1988). Hedge's g may also be computed directly from Cohen's d with the following formula: $g = d(1 - (3/(4(n_1 + n_2) - 9)))$. When there were multiple outcomes per domain they were combined according to Borenstein, Hedges, & Rothstein (2007). The overall mean effect size for all of the studies combined was computed using the following formula: $\bar{g} = \sum w_j g_j / \sum w_j$ where w_j is the weight for each study and g_j is the effect size for each study. Effect sizes were calculated with both fixed and random effects models. The fixed effects analysis estimates the exact overall effect size based on the studies included—assuming this represents the entire population of studies. The random effects analysis estimates the overall effect size assuming the studies included are only a sample of the entire population of studies. For the primary analyses (VRET versus control and VRET versus in vivo) we report both Hedge's g (to control for sample size) and Cohen's d (more common in the literature and with small, medium, and large conventions). Thereafter we only report the more conservative Hedge's g.

2. Results

Across studies, the most common control condition was no treatment/waitlist, accounting for 8 (Botella et al., in press; Difede et al., in press; Garcia-Palacios et al., 2002; Harris et al., 2002; Krijn, Emmelkamp, Biemond et al., 2004; Rothbaum et al., 1995; Rothbaum et al., 2000; Rothbaum et al., 2006) out of 11 studies (see Table 2). One study included an attention control (Maltby et al., 2002), one used bibliotherapy (Krijn et al., 2007), and one used a relaxation control (Muhlberger, Herrmann, Wiedemann, Ellgring, & Pauli, 2001). Most studies (9 of 13) targeted specific phobias ($g = 0.95$); nonetheless, the effect size of the social phobia ($g = 0.73$), PTSD ($g = 0.72$), and panic disorder ($g = 1.59$) were in the same medium–large range as the mean across all the studies.

2.1. Hypothesis 1: virtual reality exposure therapy versus control

First, using a Hedge's g random effects analysis we obtained a mean overall effect size of Hedge's $g = 1.08$ (S.E. = 0.14, 95% CI: 0.80–1.35) indicating a large effect for VRET relative to control conditions. Fig. 1 is a forest plot of the bias corrected (Cohen's d) between group (controlled) effect sizes and 95% confidence intervals for each study with virtual reality and control conditions. Using a random effects analysis we obtained a mean overall effect size of Cohen's $d = 1.11$ (S.E. = 0.15, 95% CI: 0.82–1.39), indicating a large effect for virtual reality exposure interventions relative to control conditions (see Fig. 1).

2.2. Hypothesis 2: effect sizes across dependent variables

The primary study hypotheses were tested with domain-specific measures. However, separate analyses

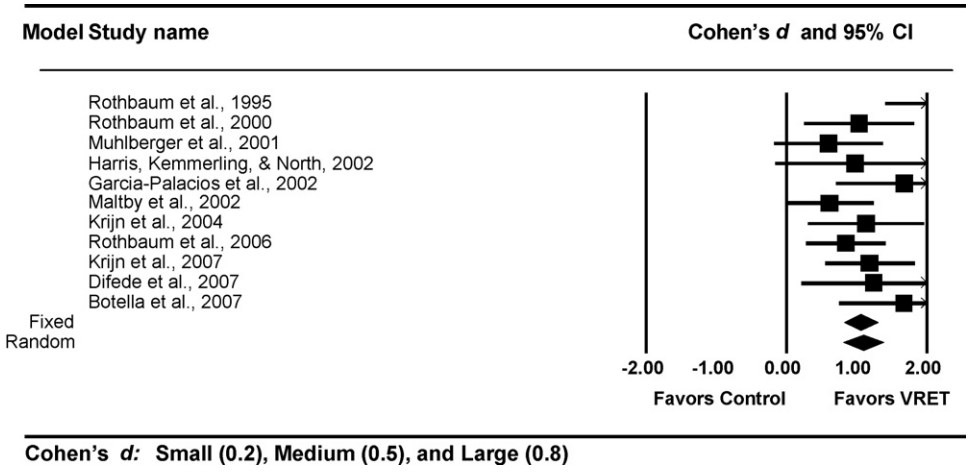

Fig. 1. Forest plot of VR vs. control meta-analysis.

Table 2
Virtual reality exposure therapy in the treatment of anxiety disorders

Study	Control	Sample	*N*	No. of sessions	Assignment	Primary outcome measures	Effect size[a]
Specific phobia							
Rothbaum et al. (1995)	Waitlist	Clinical	17	8	Random	AQ-Total	2.75
Rothbaum et al. (2000)	Waitlist	Clinical	30	8	Random	FFI and QAF	1.03
Rothbaum et al. (2000)	In vivo	Clinical	30	8	Random	FFI and QAF	−0.10
Muhlberger et al. (2001)	Relaxation	Clinical	28	1	Random	AES, FFS, AFA, and DES	0.60
Emmelkamp et al. (2002)	In vivo	Clinical	33	3	Random	AQ-anxiety and AQ-avoidance	0.07
Garcia-Palacios et al. (2002)	Waitlist	Clinical	23	4	Random	BAT SUDs and FSQ	1.68
Maltby et al. (2002)	Attention control	Clinical	43	5	Random	BAT SUDs, FAM, and FAS	0.62
Krijn, Emmelkamp, Biemond et al. (2004)	Waitlist	Clinical	28	3	Random	AQ-anxiety and AQ-avoidance	1.13
Rothbaum et al. (2006)	Waitlist	Clinical	54	4	Random	FFI and QAF	0.85
Rothbaum et al. (2006)	In vivo	Clinical	54	4	Random	FFI and QAF	0.68
Krijn et al. (2007)	Waitlist	Clinical	49	4	Random	FAS and FAM	1.19
Social phobia							
Harris et al. (2002)	Waitlist	Analogue	14	4	Random	LSAS and PRCS	0.98
Klinger et al. (2005)	In vivo	Clinical	36	12	Matched	LSAS and SCIA	0.36
PTSD							
Difede et al. (in press)	Bibliotherapy	Clinical	18	7	Matched	CAPS	1.24
Panic disorder							
Botella et al. (in press)	Waitlist	Clinical	24	9	Random	PDSS, ASI, FQ-avoidance, target fear and avoidance	1.67
Botella et al. (in press)	In vivo	Clinical	24	9	Random	PDSS, ASI, FQ-avoidance, target fear and avoidance	0.44

AES: anxiety expectancy scale, AFA: general fear of flying questionnaire (fear of flying and avoidance), AQ: acrophobia questionnaire, ASI: anxiety sensitivity index, BAT: behavioral avoidance test, CAPS: clinician administered PTSD scale, DES: danger expectancy scale, FAM: flight anxiety modality questionnaire, FAS: flight anxiety situations questionnaire, FFI: fear of flying inventory, FFS: fear of flying scale, FQ: fear questionnaire, FSQ: fear of spiders questionnaire, LSAS: Liebowitz social anxiety scale, PDSS: panic disorder severity scale, PRCS: person report of confidence as a speaker, QAF: questionnaire on attitudes toward flying, SCIA: social contexts inducing anxiety.

[a] Cohen's *d*: small (0.2), medium (0.5), and large (0.8).

were conducted for other outcome categories including general subjective distress (i.e. BDI, SCL-90, etc.), cognition, behavior (BATs), and psychophysiology. The four studies that included general distress measures (Botella et al., in press; Difede et al., in press; Harris et al., 2002; Muhlberger et al., 2001) showed a mean overall effect size of Hedge's g = 0.5 (S.E. = 0.24, 95% CI: 0.006–0.95), indicating a medium effect for virtual reality exposure interventions relative to control conditions. The five studies that included cognitive measures (Botella et al., in press; Harris et al., 2002; Krijn, Emmelkamp, Biemond et al., 2004; Maltby et al., 2002; Rothbaum et al., 1995) showed a mean overall effect size of Hedge's $g = 1.30$ (S.E. = 0.31, 95% CI: 0.70–1.91), indicating a large effect for virtual reality exposure interventions relative to control conditions. The two studies that included behavioral measures (Garcia-Palacios et al., 2002; Krijn, Emmelkamp, Biemond et al., 2004) showed a mean overall effect size of Hedge's $g = 1.27$ (S.E. = 0.31, 95% CI: 0.66–1.88), indicating a large effect for virtual reality exposure interventions relative to control conditions. Finally, the two studies that included psychophysiology measures (Harris et al., 2002; Muhlberger et al., 2001) showed a mean overall effect size of Hedge's $g = 0.68$ (S.E. = 0.33, 95% CI: 0.03–1.34), indicating a medium to large effect for virtual reality exposure interventions relative to control conditions.

2.3. Hypothesis 3: virtual reality exposure therapy versus in vivo exposure

As expected, in vivo treatment did not outperform VRET. In fact, VRET was more effective than in vivo exposure treatments. First, using a Hedge's *g* random effects analysis we obtained a mean overall effect size of Hedge's $g = 0.34$ (S.E. = 0.15, 95% CI: 0.05–0.63) indicating a small effect for VRET relative to in vivo exposure treatments. Using a Cohen's *d* random effects analysis we obtained a mean overall effect size of

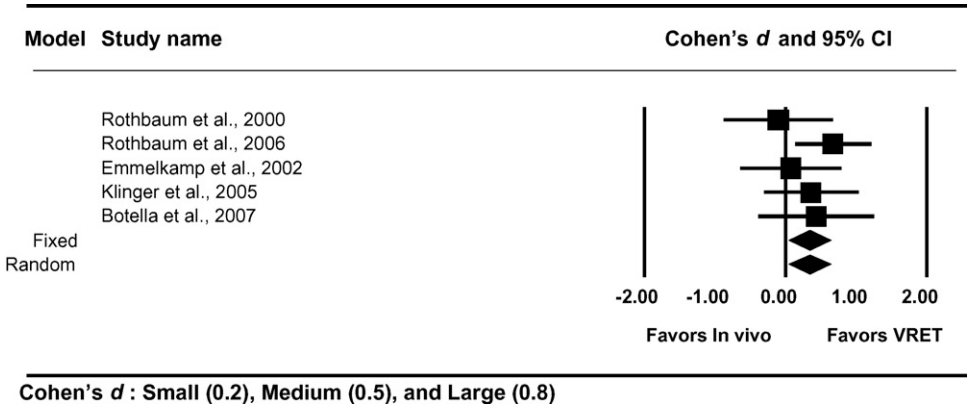

Fig. 2. Forest plot of VR vs. in vivo meta-analysis.

Cohen's $d = 0.35$ (S.E. = 0.15, 95% CI: 0.02–0.65), indicating a small effect for virtual reality exposure interventions relative to in vivo (see Fig. 2). Fig. 2 is a forest plot of the bias corrected (Cohen's d) between group (controlled) effect sizes and 95% confidence intervals for each study with a virtual reality and in vivo condition.

The degree of match between VRET and in vivo exposure environments varied across studies. The study on acrophobia matched the in vivo and VRET exposure environments exactly including a multilevel mall, a fire escape, and a rooftop garden (Emmelkamp et al., 2002). The two studies of flight fear mostly matched exposure environments such as waiting on the plane (Rothbaum et al., 2000, 2006). The studies on social phobia (Klinger et al., 2005) and panic disorder (Botella et al., in press) were both conducted within a standard CBT format but varied the exposure delivery method (in vivo or VRET).

2.4. Hypothesis 4: dose–response relationship

A meta-regression analysis showed a trend for a dose–response relationship ($\beta = 0.11$, $p = 0.06$) with more treatment sessions yielding larger effect sizes (see Fig. 3).

2.5. Hypothesis 5: effect size as a function of sample size and publication year

A meta-regression analysis showed there was no significant relationship between sample size and effect size ($\beta = -0.02$, $p = 0.10$). Likewise, there was no significant relationship between publication year and effect size ($\beta = -0.02$, $p = 0.70$).

2.6. Publication bias "the file drawer problem"

Several authors suggest there may be a potential discrepancy between the number of published trials and

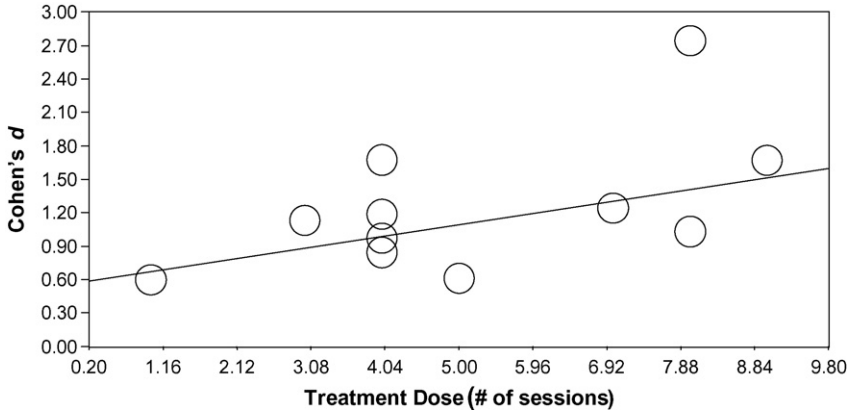

Fig. 3. Dose–response regression.

the total number that are completed (Bakan, 1967; McNemar, 1960; Smart, 1964; Sterling, 1959). In this way any meta-analysis of published studies may be missing non-significant findings and therefore overestimate the overall effect size. Rosenthal (1991) and others have called this confound "the file drawer problem". A conservative method of addressing this problem is to assume that the effect sizes of all current or future *unpublished* studies are equal to 0 and compute the number of such studies it would require to reduce the overall effect size to a non-significant level (Rosenthal & Rubin, 1988). This value may be referred to as the "fail-safe N".

Rosenthal (1991) suggested the following equation to compute a fail-safe N: $X = (K(K\overline{Z}^2 - 2.706))/2.706$ where K is the number of studies in the meta-analysis and $\overline{Z}$ is the mean Z obtained from the K studies. Rosenthal (1991) suggested that findings may be considered robust if the required number of studies (X) to reduce the overall effect size to a non-significant level exceeded $5K + 10$ which in this study would be 75. Analyses revealed that it would require more than 231 current or future unpublished studies with an effect size of 0 to bring the overall effect size of the primary analysis within the non-significant range, suggesting that the meta-analysis is robust.

3. Discussion

3.1. Major findings

Our meta-analysis of 13 ($n = 397$) virtual reality exposure therapy (VRET) studies largely supported the study hypotheses. Consistent with prediction VRET showed a large overall effect size compared to control conditions, Cohen's $d = 1.11$ (S.E. = 0.15, 95% CI: 0.82–1.39). This result was consistent across secondary outcome variables (general subjective distress, cognition, behavior, and psychophysiology). As predicted, in vivo exposure was not significantly more effective than VRET. In fact, VRET outperformed in vivo exposure, Cohen's $d = 0.35$ (S.E. = 0.15, 95% CI: 0.05–0.65). This was particularly interesting given most studies comparing VRET and in vivo exposure showed no such advantage taken alone. Contrary to prediction the dose–response relationship did not reach significance. However, there was a trend for more sessions yielding larger effect sizes ($p = 0.06$). Consistent with prediction neither sample size nor publication year was associated with effect sizes.

It was interesting to note the superiority of VRET over in vivo exposure when combining these studies. Several explanations are possible for the superiority of VRET over in vivo exposure. First, there may have been a weak response to in vivo conditions in these studies. However, a post hoc analysis suggested this was not due to poor performance of in vivo conditions. Although there were only three studies that directly compared in vivo treatment to a control condition, in vivo treatment showed a large effect size relative to control conditions, Cohen's $d = 1.65$ (S.E. = .78, 95% CI: 0.12–3.17). Second, credibility and expectancy may have been higher for VRET compared to in vivo exposure. Studies show a significant positive relationship between outcome and patient rated credibility and expectancy in any given condition (Jacobson & Baucom, 1977; Kazdin, 1979; Kazdin & Krouse, 1983). Future studies may benefit from including data on credibility and expectancy ratings between VRET and in vivo conditions to test this hypothesis (Devilly & Borkovec, 2000). Third, patients may have progressed through their hierarchy more rapidly in the VRET conditions due to a perception of increased control and safety. Of note, a gradual linear increase in self-efficacy ratings is found across VRET sessions (Krijn et al., 2007, in press). Additional possible reasons include: (a) allegiance effects for VRET, (b) exposures could be more personally tailored by the experimenter in the VRET conditions, and (c) outcome measures favored VRET. This last point would be addressed by inclusion of more behavioral measures in future trials.

A number of studies combined VR treatment with cognitive techniques (Krijn et al., in press; Muhlberger et al., 2003) and were excluded from the meta-analysis. Results of cognitive restructuring in these studies are difficult to evaluate since an amalgam of different cognitive techniques was used. To date, only one study (Krijn et al., in press) investigated whether the addition of cognitive self-statements to VRET enhanced effectiveness. Results were negative: the combined cognitive-VRET treatment was not more effective than VRET alone. However, it is interesting to note that in this meta-analysis VRET showed a very large effect size ($g = 1.30$) for cognitive outcome measures. There is a clear need for more detailed analysis of the role of cognitive techniques in VR exposure.

3.2. Presence

"Presence" is the extent to which virtual reality devices feel realistic to participants (Price & Anderson, in press). If VRET reduces fear by activating a fear structure (Foa & Kozak, 1986) then it is assumed that the more VRET "feels" realistic the better it should

work. Unfortunately, very few studies of VRET have assessed 'presence', thus precluding a meta-analysis of these data. Only one study to date has experimentally examined multiple levels of immersion in VRET. Krijn, Emmelkamp, Biemond et al. (2004), Krijn, Emmelkamp, Olafsson et al. (2004) randomized participants to either a control condition or one of two levels of immersion: (a) VRET with a head-mounted display (HMD) or (b) a computer automatic virtual environment (CAVE) (Krijn, Emmelkamp, Biemond et al., 2004). Results showed that both treatments were equally effective and superior to no treatment. It is therefore likely that some level of presence is necessary for VRET to be effective. However, thus far efforts to increase presence have not resulted in superior outcome.

3.3. Limitations

Although a few studies have also investigated VRET in panic disorder, PTSD, and social phobia, most studies have been conducted with specific phobias, most notably fear of flying and acrophobia. Therefore, generalization of the results of the present meta-analysis should be interpreted with caution with respect to other anxiety disorders. Nevertheless, effect sizes achieved in these other disorders were more or less of the same magnitude as those achieved in specific phobias. Another limitation is that very few studies have used a behavioral measure. The studies which did so, however, found large effect sizes. Future research would profit from the addition of a behavioral avoidance test to analyze the effect of treatment on real anxiety provoking situations and generalization to the real world.

3.4. Summary and conclusion

In sum, VRET is highly effective in treating phobias and more so than inactive (waiting list and attention control) and active (relaxation and bibliotherapy) control conditions. Interestingly, this meta-analysis revealed that VRET is slightly but significantly more effective than exposure in vivo, the gold standard in the field. There are a number of advantages of VRET over exposure therapy (Emmelkamp, 2005). The treatment can be conducted in the therapist's office rather than the therapist and patient having to go outside to do the exposure exercises in real phobic situations. Further, VRET provides the possibility of generating more gradual assignments (sequence and intensity of treatment), and of creating idiosyncratic exposure. In the treatment of fear of flying, the advantages of VRET over standard exposure therapy are enormous. It is highly cost effective, components of the flight can be repeated endlessly in the therapist office, and different flight destinations, different crews, and different weather conditions can be created in seconds. Another advantage is that VR treatment can also be applied to patients who are too anxious to undergo real-life exposure in vivo. Given these advantages and the efficacy of VRET supported by this meta-analysis a broader application in clinical practice seems justified.

References

Bakan, D. (1967). *On method.* San Francisco, CA: Josey-Bass.

Borenstein, M., & Rothstein, H. (1999). *Comprehensive meta-analysis a computer program for research synthesis.* New Jersey: Biostat Inc.

Borenstein, M., Hedges, L. V., & Rothstein, H. (2007). *Introduction to meta-analysis.* Chichester, England: Wiley.

Botella, C., Villa, H., Garcia-Palacios, A., Banos, R. M., Quero, S., Alcaniz, M., et al. (in press). Virtual reality exposure in the treatment of panic disorder and agoraphobia: A controlled study. *Clinical Psychology & Psychotherapy.*

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Deacon, B. J., & Abramowitz, J. S. (2004). Cognitive and behavioral treatments for anxiety disorders: A review of meta-analytic findings. *Journal of Clinical Psychology, 60*(4), 429–441.

Devilly, G. J., & Borkovec, T. D. (2000). Psychometric properties of the credibility/expectancy questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry, 31*(2), 73–86.

Difede, J., Cukor, J., Jayasinghe, N., Patt, I., Jedel, S., & Spielman, L. (in press). Virtual reality therapy for the treatment of posttraumatic stress disorder following September 11, 2001. *Journal of Clinical Psychiatry.*

Eddy, K. T., Dutra, L., Bradley, R., & Westen, D. (2004). A multidimensional meta-analysis of psychotherapy and pharmacotherapy for obsessive-compulsive disorder. *Clinical Psychology Review, 24*(8), 1011–1030.

Emmelkamp, P. M. (2005). Technological innovations in clinical assessment and psychotherapy. *Psychotherapy and Psychosomatics, 74*(6), 336–343.

Emmelkamp, P. M., Bryunzeel, M., Drost, L., & van der Mast, C. A. P. G. (2001). Virtual reality treatment in acrophobia: A comparison with exposure in vivo. *Cyberpsychology and Behavior, 4*, 335–345.

Emmelkamp, P. M., Krijn, M., Hulsbosch, A. M., de Vries, S., Schuemie, M. J., & van der Mast, C. A. (2002). Virtual reality treatment versus exposure in vivo: A comparative evaluation in acrophobia. *Behavior Research and Therapy, 40*(5), 509–516.

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin, 99*(1), 20–35.

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*(9), 983–993.

Garcia-Palacios, A., Hoffman, H. G., See, S. K., Tsai, A., & Botella, C. (2001). Redefining therapeutic success with virtual reality exposure therapy. *Cyberpsychology and Behavior, 4*(3), 341–348.

Gould, R. A., Otto, M. W., & Pollack, M. H. (1995). A meta-analysis of treatment outcome for panic disorder. *Clinical Psychology Review, 15*(8), 819–844.

Gould, R. A., Otto, M. W., Pollack, M. H., & Yap, L. (1997). Cognitive behavioral and pharmacological treatment of generalized anxiety disorder: A preliminary meta-analysis. *Behavior Therapy, 28*(2), 285–305.

Harris, S. R., Kemmerling, R. L., & North, M. M. (2002). Brief virtual reality therapy for public speaking anxiety. *Cyberpsychology and Behavior, 5*(6), 543–550.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Olrando, FL: Academic Press.

Jacobson, N. S., & Baucom, D. H. (1977). Design and assessment of nonspecific control groups in behavior modification research. *Behavior Therapy, 8*(4), 709–719.

Kazdin, A. E. (1979). Therapy outcome questions requiring control of credibility and treatment-generated expectancies. *Behavior Therapy, 10*(1), 81–93.

Kazdin, A. E., & Krouse, R. (1983). The impact of variations in treatment rationales on expectancies for therapeutic change. *Behavior Therapy, 14*(5), 657–671.

Klinger, E., Bouchard, S., Legeron, P., Roy, S., Lauer, F., Chemin, I., et al. (2005). Virtual reality therapy versus cognitive behavior therapy for social phobia: A preliminary controlled study. *CyberPsychology & Behavior, 8*(1), 76–88.

Kobak, K. A., Greist, J. H., Jefferson, J. W., Katzelnick, D. J., & Henk, H. J. (1998). Behavioral versus pharmacological treatments of obsessive compulsive disorder: A meta-analysis. *Psychopharmacology, 136*(3), 205.

Krijn, M., Emmelkamp, P. M. G., Biemond, R., de Wilde de Ligny, C., Schuemie, M. J., & van der Mast, C. A. P. G. (2004). Treatment of acrophobia in virtual reality: The role of immersion and presence. *Behavior Research and Therapy, 42*(2), 229–239.

Krijn, M., Emmelkamp, P. M. G., Olafsson, R. P., & Biemond, R. (2004). Virtual reality exposure therapy of anxiety disorders: A review. *Clinical Psychology Review, 24*(3), 259–281.

Krijn, M., Emmelkamp, P. M., Olafsson, R. P., Bouwman, L. J., Van Gerwen, L. J., Spinhoven, P. H., et al. (2007). Fear of flying treatment methods: Virtual reality exposure vs. cognitive behavioral therapy. *Aviation, Space and Environmental Medicine, 78*, 121–128.

Krijn, M., Emmelkamp, P. M., Olafsson, R. P., Schuemie, M. J., & van der Mast, C. A. (in press). Do self-statements enhance the effectiveness of virtual reality exposure therapy? A comparitive evaluation in acrophobia. *CyberPsychology & Behavior.*

Maltby, N., Kirsch, I., Mayers, M., & Allen, G. J. (2002). Virtual reality exposure therapy for the treatment of fear of flying: A controlled investigation. *Journal of Consulting and Clinical Psychology, 70*(5), 1112–1118.

McNemar, Q. (1960). At random: Sense and nonsense. *American Psychologist, 15*, 295–300.

Muhlberger, A., Herrmann, M. J., Wiedemann, G., Ellgring, H., & Pauli, P. (2001). Repeated exposure of flight phobics to flights in virtual reality. *Behaviour Research and Therapy, 39*(9), 1033–1050.

Muhlberger, A., Wiedemann, G., & Pauli, P. (2003). Efficacy of a one-session virtual reality exposure treatment for fear of flying. *Psychotherapy Research, 13*(3), 323–336.

North, M. M., North, S. M., & Coble, J. R. (1998). Virtual reality therapy: An effective treatment for the fear of public speaking. *International Journal of Virtual Therapy, 3*, 2–8.

Price, M., & Anderson, P. (in press). The role of presence in virtual reality exposure therapy. *Journal of Anxiety Disorders*.

Ressler, K. J., Rothbaum, B. O., Tannenbaum, L., Anderson, P., Graap, K., Zimand, E., et al. (2004). Cognitive enhancers as adjuncts to psychotherapy: Use of D-cycloserine in phobic individuals to facilitate extinction of fear. *Archives of General Psychiatry, 61*(11), 1136–1144.

Rosenthal, R. (1991). *Meta-analytic procedures for social research*. London, UK: Sage.

Rosenthal, R., & Rubin, D. B. (1988). Comment: Assumptions and procedures in the file drawer problem. *Statistical Science, 3*, 120–125.

Rothbaum, B. O., Anderson, P., Zimand, E., Hodges, L., Lang, D., & Wilson, J. (2006). Virtual reality exposure therapy and standard (in vivo) exposure therapy in the treatment of fear of flying. *Behavior Therapy, 37*(1), 80–90.

Rothbaum, B. O., Hodges, L. F., Kooper, R., Opdyke, D., Williford, J. S., & North, M. (1995). Effectiveness of computer-generated (virtual reality) graded exposure in the treatment of acrophobia. *American Journal of Psychiatry, 152*(4), 626–628.

Rothbaum, B. O., Hodges, L., Smith, S., Lee, J. H., & Price, L. (2000). A controlled study of virtual reality exposure therapy for the fear of flying. *Journal of Consulting and Clinical Psychology, 68*(6), 1020–1026.

Smart, R. G. (1964). The importance of negative results in psychological research. *Canadian Psychologist, 5A*, 225–232.

Sterling, T. D. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance—or vice versa. *Journal of the American Statistical Association, 54*, 30–40.

Vincelli, F., Anolli, L., Bouchard, S., Wiederhold, B. K., Zurloni, V., & Riva, G. (2003). Experiential cognitive therapy in the treatment of panic disorders with agoraphobia: A controlled study. *CyberPsychology & Behavior, 6*(3), 321–328.

Wiederhold, B. K., Jang, D. P., Gevirtz, R. G., Kim, S. I., Kim, I. Y., & Wiederhold, M. D. (2002). The treatment of fear of flying: A controlled study of imaginal and virtual reality graded exposure therapy. *IEEE Transactions on Information Technology in Biomedicine, 6*(3), 218–223.