Visuo-spatial perspective-taking, avatar embodiment and the ability to cultivate compassion using virtual reality and mental imagery

Nicola Alden

D.Clin.Psy. thesis (Volume 1), 2015

University College London

UCL Doctorate in Clinical Psychology

Thesis declaration form

I confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature:

Name: Nicola Alden

Date: 19.06.2015

Overview

The use of immersive virtual reality (IVR) as a tool for treating psychological difficulties is a rapidly developing field. Presence, the sense of being in a virtual environment, is regarded as a key mechanism underlying the efficacy of treatment using IVR. Part 1 of this thesis is a systematic review and meta-analysis of the relationship between presence and the outcome of treatment using IVR with clinical populations. Variables which potentially moderate this relationship were examined using narrative synthesis. Greater presence during IVR was associated with a more positive treatment outcome but this association may be moderated by treatment type.

Part 2, an empirical study, examined the relationship between allocentric visuo-spatial perspective-taking ability, sense of embodiment and the ability to cultivate self-compassion in self-critical individuals using an IVR or analogue mental imagery intervention. Change in self-compassion and self-criticism following the intervention was examined. Experience of the intervention and effects related to practicing imagining the intervention for two weeks were assessed. Self-criticism reduced after both interventions but self-compassion increased only after the mental imagery intervention. Rather than visuo-spatial perspective-taking ability or embodiment it seems the experience of the intervention may have contributed to these findings. This study was part of a joint project (see Holden, n.d.).

Part 3 is a critical appraisal of parts 1 and 2. It reflects on processes involved in carrying out part 1 then expands on methodological choices made in part 2. Finally, issues with using IVR in the empirical study and therapy more generally are discussed.

Table of Contents

Table of Tables

Part 1: Literature Review

Part 2: Empirical Study

Table of Figures

Part 1: Literature Review

Part 2: Empirical Study

Acknowledgments

I would like to thank my supervisor, Dr John King, for his kindness and enthusiasm throughout this project and his thoughtful and nurturing approach to supervision. Thanks also to my caring and supportive partner for the empirical study, Anneka Holden, whom it was an absolute pleasure to work with.

I am grateful for help from Dr Caroline Falconer who gave her time to teach me how to use the immersive virtual reality technology and talk through aspects of the empirical study. I am also grateful to Aitor Rovira whose expertise in immersive virtual reality helped the experimental sessions to run smoothly. I wish to thank those who participated in the empirical study for their generosity with their time and interest in the project.

I would especially like to thank my family for their love and support.

Part 1: Literature Review

The relationship between presence and the outcome of treatment using immersive virtual reality: A systematic review and meta-analysis

Abstract

Aims

Presence, the sense of being in a virtual environment, may be a key mechanism underlying the efficacy of treatment using immersive virtual reality (IVR). This review assessed the relationship between presence and the outcome of psychological treatment using IVR and examined variables that possibly moderate this relationship.

Method

Searches for studies assessing this relationship were conducted using PsycINFO, PubMed and Web of Science databases and manual searching. Studies were identified against pre-specified criteria. Their methodological quality was assessed. Meta-analysis was conducted using Comprehensive Meta-Analysis, Version 3.2. Results of studies ineligible for meta-analysis were summarised. Possible moderating variables were examined using narrative synthesis.

Results

Twelve studies met the criteria. Seven treated anxiety using exposure or relaxation IVR treatment. Four used IVR as a distraction tool to treat pain or symptom distress. One used IVR tasks to treat tobacco addiction. Meta-analysis included seven studies (n = 311). This showed a small effect for the relationship between presence and the outcome of treatment using IVR. Two studies ineligible for meta-analysis indicated this relationship. Probable moderating variables included treatment type.

Conclusions

Greater presence during treatment using IVR is associated with a more positive treatment outcome but this association may be moderated by treatment type. Adequately powered, controlled studies are needed for robust conclusions along with further examination of moderating factors.

1: Introduction

1.1 Psychological Treatment Using Immersive Virtual Reality

Immersive virtual reality (IVR) integrates sensory input, interactive programmes, body tracking devices and real-time computer graphics to immerse an individual in a computer-generated virtual environment (VE). The VE is usually presented through a head-mounted display, a helmet with headphones and visual screens. Alternatively, a computer automated virtual environment system can be used. This projects the VE on the floors and walls instead of using a helmet. Given that customised VEs can be created, IVR technology offers a new tool for psychological therapy. Indeed, the use of IVR in psychotherapy is a rapidly developing field (Riva, 2005) with therapeutic applications being developed for a wide range of psychological treatments (Gregg & Tarrier, 2007).

1.1.1 Anxiety disorders

One line of investigation has examined IVR as a tool for providing exposure therapy. Exposure therapy, the most effective evidence-based intervention for anxiety disorders, is traditionally carried out using in vivo or imaginal exposure (Parsons & Rizzo, 2008). However many individuals are often reticent to seek out or fully engage with such treatments (Powers & Emmelkamp, 2008). Compared to in vivo exposure, virtual reality exposure therapy (VRET) may be more convenient, confidential and cost-effective. Additionally, VRET offers greater control over exposure and thus over the level of arousal making it less prone to flooding (Gerardi, Cukor, Difede, Rizzo, & Rothbaum, 2010). This may make VRET more acceptable than its alternatives. In their narrative review Meyerbröker and Emmelkamp (2010) concluded that VRET may be an effective method for delivering exposure treatment

for anxiety disorders. Recent meta-analyses reached similar conclusions (McCann et al., 2014; Opriş et al., 2012; Parsons & Rizzo, 2008; Powers & Emmelkamp, 2008). Overall they found that VRET was superior to no treatment and equipotent to active comparison groups which used traditional forms of exposure therapy.

IVR has also been investigated as a relaxation tool for stress treatment. Relaxation in IVR is facilitated by positive visual and auditory stimulation and contextual cues. Research has found IVR to be comparable to DVD and audio tape treatments for stress (Villani & Riva, 2008). However, using IVR as a therapeutic relaxation tool offers convenience and control. Furthermore, it enables the individual to live the experience in a vivid and realistic manner. This removes reliance on the individual's capacity to generate images, as required by many traditional relaxation techniques (Vincelli, 1999).

1.1.2 Tobacco addiction

The ability of IVR to provide ecologically valid contexts and controlled, individualised treatment programmes has also led to interest in using this tool for cue-exposure therapy. Cue-exposure therapy is commonly used to treat addiction. It involves the repeated presentation of addiction related cues with the aim of extinguishing cue reactivity. This is typically carried out in laboratory settings using two-dimensional stimuli. However, its effectiveness is limited (Bordnick, Carter, & Traylor, 2011). Encouragingly, initial research investigating treatment for tobacco addiction using IVR-cue-exposure therapy (Moon & Lee, 2009; Park et al., 2014) and related IVR paradigms such as crushing virtual cigarettes (Girard, Turcotte, Bouchard, & Girard, 2009) shows promise.

1.1.3 Pain/symptom distress

Another area of research has investigated IVR as an alternative to traditional distraction techniques for reducing pain associated with physical health difficulties (Wismeijer & Vingerhoets, 2005). Pain exerts a powerful demand for attentional resources (Melzack & Wall, 1965). Any task or stimulus that competes for these resources leaves less cognitive capacity available for processing pain. Thus, IVR distraction could function as a non-pharmacological form of analgesia (Mahrer & Gold, 2009; Morris, Louw, & Grimmer-Somers, 2009). A small number of studies have shown that IVR distraction is effective for reducing pain and other unpleasant symptoms associated with medical procedures both as a standalone and adjunctive intervention (Mahrer & Gold, 2009; Morris et al., 2009; Wismeijer & Vingerhoets, 2005). The effectiveness of this pain management tool has been found to increase with greater levels of interactivity and "fun" (Malloy & Milling, 2010; Triberti, Repetto, & Riva, 2014; Wiederhold & Wiederhold, 2007).

1.1.4 Limitations

Despite the promising results, investigations of treatment using IVR must be regarded with some caution due to concerns common to newly developing areas of research such as methodological rigour (McCann et al., 2014). For example, reviews have often combined evidence from clinical and non-clinical populations. However, significant differences have been found between the two groups in their response to IVR interventions (Ling, Nefs, Morina, Heynderickx, & Brinkman, 2014). Furthermore, the underlying mechanisms and factors which moderate the effectiveness of treatment using IVR remain unclear (Parsons & Rizzo, 2008).

1.2 Presence

Presence in virtual reality has commonly been defined as the psychological sense of "being there". It is the subjective experience of being in the VE rather than the environment you are physically located in (Witmer & Singer, 1998). Numerous theories of the nature of presence have been proposed most of which suggest that this construct is derived from several factors. Across different theoretical constructs these factors commonly include the individual's sense of spatial presence (being spatially located in the VE), involvement with the VE, control over the VE and the realness of the VE (Schuemie, van der Straaten, Krijn, & van der Mast, 2001). Different measures of presence have been developed based on these different theoretical constructs. It is worth noting that a distinction is made between the concepts of presence and immersion. Immersion is an objective description of the technical capability of the IVR technology to generate a surrounding, convincing and interactive VE (Sanchez-Vives & Slater, 2005; Slater & Wilbur, 1997). Presence is the individual's subjective response to the VE. Therefore, presence may be influenced by the individual's state of mind (Bowman & McMahan, 2007) and how effectively they are immersed in the VE (Hoffman et al., 2004).

1.3 The Relationship Between Presence and the Outcome of Treatment Using IVR

Presence is regarded as a key mechanism underlying the efficacy of treatment using IVR. It has been assumed that the more the individual feels present in the VE when the treatment is administered the more deeply experienced it will be with positive implications for the treatment outcome (Powers & Emmelkamp, 2008; Spagnolli, Bracken, & Orso, 2014; Triberti et al., 2014; Wismeijer & Vingerhoets,

2005). Recent systematic reviews have examined the role of presence in treatment using virtual reality. Spagnolli et al. (2014) reviewed the role of presence in validating the efficacy of cyber-therapy (i.e., psychological therapy carried out using any type of mediated environment including computers, mobile phones or TV). They argued that sense of presence is interconnected with cyber-therapy and therefore plays a useful role in the validation of cyber-therapy treatment. Triberti et al. (2014) combined data from clinical and non-clinical populations to review the use of both immersive and non-immersive virtual reality distraction interventions for pain. Their findings suggest that sense of presence influences the effectiveness of this type of treatment.

Given that the outcome of treatment using IVR and non-immersive virtual reality differs (Malloy & Milling, 2010), these reviews do not address the role of presence in treatment specifically using IVR. Additionally, the response of clinical populations to IVR treatment remains unclear (Ling et al., 2014). Thus, this review systematically examined empirical investigations of the relationship between presence and the outcome of treatment using IVR in clinical populations. As the sense of presence experienced in IVR is typically assessed using self-report questionnaires the review focused on studies using this type of measurement. Such investigations report mixed results (e.g., see conflicting results reported by Hoffman et al., 2008, and Price & Anderson, 2007). However, these investigations vary in participant, technology and design characteristics which may influence their findings (Hoffman et al., 2004; Meyerbröker & Emmelkamp, 2010; Triberti et al., 2014). Therefore, such variables require consideration in order to effectively interpret the results. Furthering understanding of the relationship between presence and the outcome of treatment using IVR will help with the development of effective IVR

treatments. Moreover, it will help implement this tool into clinical practice by developing the evidence on predicting treatment outcomes (Malloy & Milling, 2010; Triberti et al., 2014). This is important given that individuals with psychological difficulties and health care providers may find IVR preferable to alternative forms of treatment (Gorini & Riva, 2008; Wiederhold & Wiederhold, 2007).

1.4 Aims

In sum, the primary aim of this review is to assess the relationship between presence and the outcome of treatment using IVR in clinical populations. The secondary aim is to assess variables which may influence this relationship. Based on the existing literature the following variables are investigated: participant demographics, clinical difficulty, technology characteristics and study design characteristics (i.e., measures, type of treatment using IVR, dose-response relationship and sample size). Clear hypotheses about the direction of effects are not made due to mixed results in the literature.

The methodological quality of the included studies is assessed. Following this the primary aim of the review is investigated using meta-analysis. The results of studies which were ineligible for inclusion in the meta-analysis are summarised. The secondary aim is addressed using narrative synthesis.

2: Method

2.1 Criteria for Considering Studies for this Review

2.1.1 Types of studies

Only studies using IVR as a treatment tool were considered eligible for inclusion. Any therapeutic application of IVR was acceptable and the treatment

protocol could include multiple components (e.g., CBT with IVR used for exposure treatment). Studies using non-immersive virtual reality (e.g., augmented reality, computer simulations or mobile phones), analogue treatments (e.g., experimentally induced pain) and investigations of dimensions of the participant's experience where this was not in and of itself a treatment (e.g., the ability of IVR to elicit anger) were excluded.

To be eligible studies were required to statistically analyse the effect of the sense of presence experienced by the participant during treatment using IVR on the outcome of this treatment.

Only studies published in peer-reviewed journals and in English were included. Case series, case studies, conference proceedings, qualitative studies and studies which did not report original empirical findings were excluded.

2.1.2 Types of participants

Any clinical population was acceptable. No restrictions were placed on participant demographics or the setting for recruitment or treatment using IVR.

2.1.3 Types of outcome measures

Studies were included if they used a quantitative self-report measure of the sense of presence experienced by the participant during treatment using IVR and quantitative self-report measure(s) of the difficulty this treatment was designed to address.

2.2 Search Methods for Identification of Studies

2.2.1 Electronic searches

Systematic searches were run in PsycINFO, PubMed and Web of Science databases for studies published all years up until August 2014. The results were limited to journal articles and articles published in English.

Several search strategies were trialled in order to develop a strategy which maximised the sensitivity and precision of the search. Initially, subject heading terms were identified and included in searches run in PsycINFO and PubMed (e.g., human-computer interaction, user-computer interface). These terms proved either over-inclusive in the types of records returned or failed to return any relevant records over and above keyword searches. Initially included keyword terms such as virtual environment also failed to return relevant records. Therefore, these terms were removed from the search strategy.

The final search used the keywords virtual reality and presence. Terms were entered individually and were then combined, that is, (virtual reality) AND presence.

2.2.2 Searching other resources

Manual searches were conducted on the reference lists of the studies included in the review and 15 published systematic reviews and meta-analyses related to the review topic. Publication lists of key researchers in relevant fields were also searched manually.

2.3 Data Collection and Analysis

2.3.1 Selection of studies

Study selection was carried out by the review author. Following de-duplication of the search results all identified titles and abstracts were inspected. Full reports of potentially eligible abstracts were obtained and examined to identify studies for inclusion.

2.3.2 Data extraction

The following data were extracted from the included studies:

- Study information: authors, publication year, country
- Participants' characteristics: age, gender, percentage Caucasian (the most commonly reported ethnicity data), clinical difficulty
- Sample size: for the overall study, for treatment using IVR
- Treatment using IVR: treatment type, description, duration, frequency
- IVR technology characteristics: display type, stereoscopy, tracker, interaction with the VE (characteristics examined were limited to the data available in the included studies)
- Measures (mentioned above)
- Methodological quality (see below)

Corresponding authors for the studies were contacted to obtain unclear or missing information. Two authors of one study (Dr P. Anderson and Dr M. Price) supplied information.

2.3.3 Assessment of methodological quality

Criteria for assessing methodological quality were based on established guidelines (Downs & Black, 1998; Higgins & Green, 2011). Selection was determined by the aims of the review and accounted for the mix of designs employed across the studies. Criteria were:

- Reporting: aims and/or hypotheses, inclusion/exclusion criteria, demographics, clinical presentation, confounds/adverse events, measures, administration of measures, IVR intervention described, IVR frequency and duration, technology characteristics, means and random variability for measures, statistical test, actual probability value
- External validity: representativeness of those asked to participate and the number prepared to participate, administration and setting, affordable/available technology
- Internal validity – bias: data dredging, appropriate statistical tests, compliance with the intervention, accuracy of outcome measures
- Internal validity – confounding: different groups recruited from the same population at the same time (where applicable), randomisation, allocation concealment, adjustment for confounds in the analysis, attrition accounted for
- Power
- Funding

2.3.4 Data synthesis

Statistical data relating to the relationship between presence and the outcome of treatment using IVR was synthesised using meta-analysis (described below). The results of studies which were ineligible for meta-analysis were summarised.

Sub-group analysis and meta-regression were considered to examine potentially moderating variables. However, it was decided that interpretation of the analyses would not be meaningful due to the small number of included studies, their clinical and methodological heterogeneity and their small sample sizes (Higgins & Green, 2011; Sun, Briel, Walter, & Guyatt, 2010). Therefore, narrative synthesis was carried out.

2.3.4.1 Measurement of overall effect size

Meta-analysis was completed with the software Comprehensive Meta-Analysis, Version 3.2 (Borenstein, Hedges, Higgins, & Rothstein, 2014). As the majority of studies used correlational analysis the correlation coefficient (*r*) was chosen as the effect size index. Effect size was interpreted as small *r* = .1, medium *r* = .3, or large *r* = .5 (Cohen, 1988). Four studies used correlational analysis, three used multiple regression. The correlation for one study using multiple regression (Price, Mehta, Tone, & Anderson, 2011) was provided via correspondence. Girard et al. (2009) only reported the semi-partial correlation. Therefore this was entered into the analysis. For the other study (Price & Anderson, 2007), the relevant regression coefficient was converted to a semi-partial correlation using the formula:

$$r_{sp} = \frac{t_f\sqrt{\left(1 - R_y^2\right)}}{\sqrt{(n - p - 1)}}$$

Where r_{sp} is the semi-partial correlation, t_f is the *t* test of the regression coefficient beta in the multiple regression model, R_y^2 is the total variance accounted for by the full model, n is the number of cases and p is the number of predictors (Aloe, 2014).

The sampling distribution of a correlation coefficient depends on the strength of the correlation (Borenstein, Hedges, Higgins, & Rothstein, 2009). Therefore, to enable analysis of different coefficients, the coefficients were transformed to normally distributed values using Fisher's Z transformation:

$$z = 0.5 \ln\left(\frac{1+r}{1-r}\right)$$

Where z is Fisher's Z, ln is the natural logarithm and r is the correlation coefficient (Corey, Dunlap, & Burke, 1998). The transformed values were used for all analyses. To present the results the summary effect and its confidence interval were changed back to correlations using the formula (Corey et al., 1998):

$$r_z = \frac{e^{2z} - 1}{e^{2z} + 1}$$

A random effects model was used to calculate the summary effect. This model allows for heterogeneity across studies by assuming that there is variation in the true effect size across different studies, and the studies included in the meta-analysis are a random sample of these different effect sizes. The summary effect is the estimate of the mean of the distribution of effects (Borenstein et al., 2009).

2.3.4.2 Unit of analysis issues

Some studies reported multiple treatment outcomes from the same sample. Where this occurred the effect sizes were transformed to Fisher's Z scores and the Z scores were averaged. The result was transformed back to *r* and this single mean effect size was entered into the analysis.

Some studies used domain-specific distress measures, some used domain-general distress measures and some used both to assess treatment outcome (see Table 1 for categorisation, guided by Powers & Emmelkamp, 2008). Different categories

of measures were considered to assess different aspects of the clinical difficulty being investigated. Therefore, categories were combined to obtain a single effect size for a study where necessary.

Table 1

Categorisation of Treatment Outcome Measures

Domain of distress	Treatment outcome measure
Domain-specific	ACQ, AQ, DEFAS, FAM, FAS, FFI, FFS, FFQ, FND, LSAS, MIA, PA/W, PDSS, PRCS, SSPS
Domain-general	3Q, 5Q, ABS II, ASI, BFNE, DASS, Faces Scale, PANAS, PFS, SAI, STAI-Y, VAS

Note.
3Q = 3 questions regarding aspects of pain experienced; 5Q = 5 questions regarding aspects of pain experienced; ABS II = Attitudes and Beliefs Scale II; ACQ = Agoraphobia Cognitions Questionnaire; AQ = Acrophobia Questionnaire; ASI = Anxiety Sensitivity Index; BFNE = Fear of Negative Evaluations Scale – Brief version; DASS = Depression Anxiety Stress Scales; DEFAS = Danger Expectations and Flying Anxiety Scale; FAM = Flight Anxiety Modality Questionnaire; FAS = Flight Anxiety Situation Questionnaire; FFI = Fear of Flying Inventory; FFS = Fear of Flying Scale; FFQ = Fear of Flying Questionnaire; FND = Fagerstrom test for Nicotine Dependence – revised; LSAS = Liebowitz Social Anxiety Scale; MIA = Mobility Inventory for Agoraphobia; PA/W = Panic Attacks per Week; PANAS = Positive And Negative Affect Scale; PDSS = Panic Disorder Severity Scale; PFS = Revised Piper Fatigue Scale; PRCS = Personal Report of Confidence as a Speaker; SAI = State Anxiety Inventory; SSPS = The Self Statements Towards Public Speaking Scale; STAI-Y = State-Trait Anxiety Inventory-Y Form; VAS = Visual Analogue Scale (to assess emotional states).

2.3.4.3 Heterogeneity

Heterogeneity was assessed by visually inspecting Forest Plots and examining *Q* and I^2 statistics. The *Q* statistic is a distributed chi-square and indicates if the distribution of effect sizes around their mean is greater than would occur from sampling error alone. This statistic has low power when the number of studies is small. Therefore, it is supplemented with the I^2 statistic. This expresses the percentage of variation that is due to heterogeneity between the studies as opposed to chance. This intuitive expression of the inconsistency between studies' results is not

dependent on the number of studies considered (Borenstein et al., 2009). I^2 was interpreted as: 0%-40% = might be unimportant, 30%-60% = may be moderate, 50%-90% = may be substantial, 75%-100% = considerable (Higgins & Green, 2011).

2.3.4.4 Publication bias

Publication bias can be assessed by visually inspecting a funnel plot (Higgins & Green, 2011). This was considered inappropriate as the analysis included less than the minimum of 10 studies recommended for valid estimation (Sterne et al., 2011).

3: Results

3.1 Description of Studies

3.1.1 Results of the search

A total of 1023 potentially relevant references were identified, 1021 through electronic searches and 2 through manual searches. The author excluded 292 duplicate references and 708 irrelevant references through screening titles and reading abstracts. The full text of 23 studies was retrieved for further assessment. Eleven of these were excluded (see Table A1, Appendix A for reasons for exclusion).

Of the 12 studies included in the review 5 could not be included in the meta-analysis due to missing or unavailable statistical data (Chan, Chung, Wong, Lien, & Yang, 2007; Hoffman, Patterson, & Carrougher, 2000; Hoffman et al., 2008; Moldovan & David, 2014; Riva, Manzoni, Villani, Gaggioli, & Molinari, 2008).

Figure 1, prepared in accordance with PRISMA guidelines (Liberati et al., 2009), summarises the study selection process.

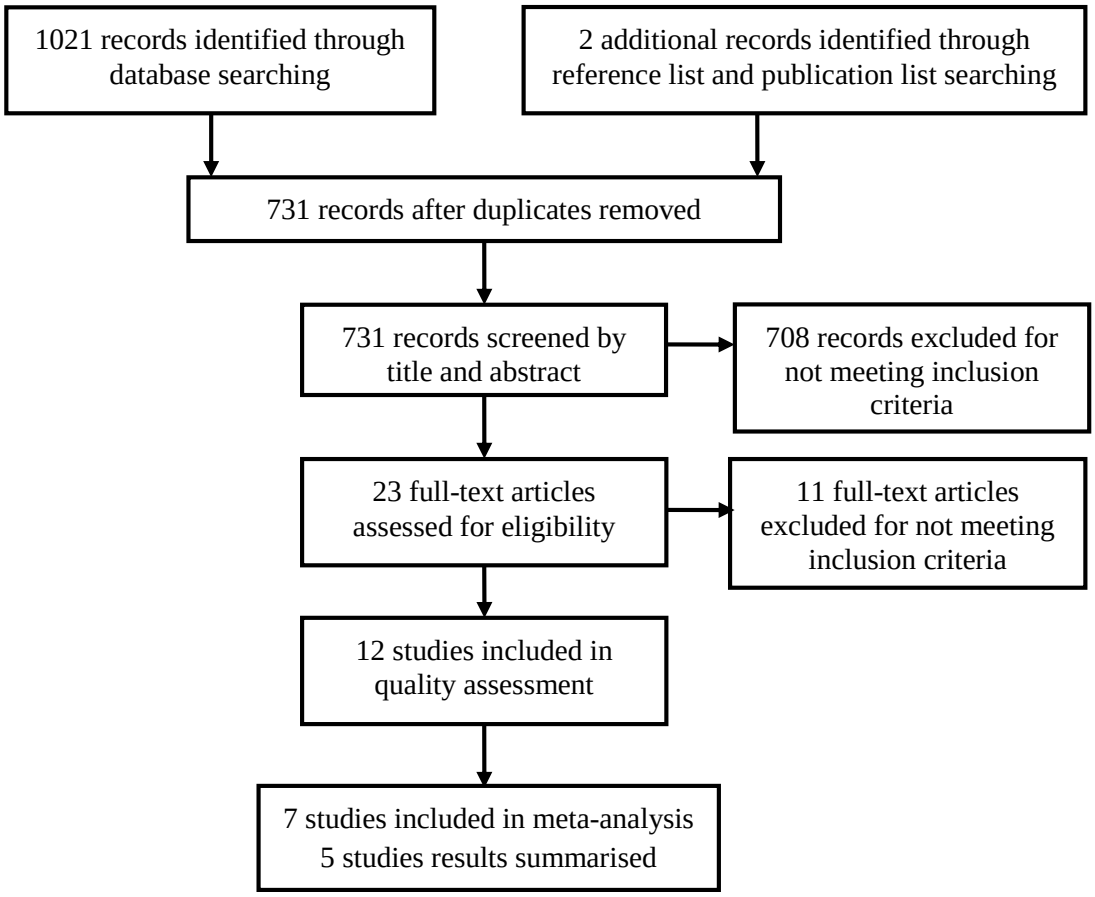


Figure 1. Study flow diagram.

3.1.2 Included studies

The following description of the included studies concerns information reported by the studies or supplied by their corresponding authors. This information is summarised in Table 2 which is ordered first by the area of clinical difficulty, then by author.

3.1.2.1 Sample size and demographics

A total of 444 participants were included in the studies, 373 of whom undertook treatment using IVR. IVR sample size ranged from 7-107 participants.

Ten studies used adult participants ranging from age 18-78 years. Hoffman et al.'s (2008) participants ranged from age 9-40 years. Chan et al.'s (2007) sample had a mean age of 7 years. The percentage of female participants ranged from 0%-100%. The percentage of Caucasian participants ranged from 54%-100%. Study populations were drawn from the following countries: Australia (1), Canada (1), Italy (1), the Netherlands (1), Spain (1), Taiwan (1), the USA (5) (unclear [1]).

3.1.2.2 Clinical difficulty

Six studies investigated a community sample with anxiety disorders. Disorders were: panic disorder with agoraphobia (Malbos, Rapee, & Kavakli, 2013; Meyerbröker, Morina, Kerkhof, & Emmelkamp, 2011), social phobia (Moldovan & David, 2014; Price et al., 2011), acrophobia (Moldovan & David, 2014), aviophobia (Moldovan & David, 2014; Rus-Calafell, Gutiérrez-Maldonado, Botella, & Baños, 2013), specific phobia or panic disorder with agoraphobia with flying as the predominantly feared stimulus (Price & Anderson, 2007). All studies confirmed participants' diagnoses using standardised DSM-IV diagnostic instruments. Across the studies common exclusion criteria included epilepsy, psychosis and substance misuse.

Riva et al. (2008) studied hospital in-patients being treated for obesity with a high level of anxiety based on healthy norms. Schneider and Hood (2007) investigated symptom distress in hospital out-patients undergoing chemotherapy. Exclusion criteria for the latter study included metastatic disease to the brain. Studies investigating pain recruited hospital in-patients in medical burns facilities (Chan et al., 2007; Hoffman et al., 2000; Hoffman et al., 2008).

Girard et al. (2009) studied a community sample of individuals who smoked 10 or more cigarettes per day. Exclusion criteria included abstinence from cigarettes for more than three months in the past year.

3.1.2.3 Treatment using IVR

The six studies investigating a community sample with anxiety disorders used VRET. Graded exposure was facilitated using a series of context-graded VEs designed to induce anxiety in the area being investigated (e.g., aviophobia - being on an aeroplane). Four studies used VEs which could be manipulated by the investigator (e.g., aviophobia - increase turbulence; Malbos et al., 2013; Price & Anderson, 2007; Price et al., 2011; Rus-Calafell et al., 2013) to enable modulation of the degree of exposure. In some studies participants stayed in each scenario either until they indicated that habituation had occurred (Price & Anderson, 2007; Rus-Calafell et al., 2013), or for a set duration during which habituation is evidenced to occur (Malbos et al., 2013). Five of these studies administered CBT techniques separately to the treatment using IVR (i.e., before or after IVR sessions). Malbos et al. (2013) randomised participants to either a CBT or a non-active intervention before all received treatment using IVR.

Riva et al. (2008) used an island VE with four zones each of which provided a different relaxing experience (e.g., waves lapping on the shore). The participant was guided through each experience by a pre-recorded therapeutic narrative. This instructed them to complete various relaxation exercises.

Girard et al. (2009) randomised participants to one of two IVR conditions. Participants explored a virtual castle to find and crush either virtual cigarettes (experimental condition) or virtual balls (control). The therapeutic task was

administered in conjunction with the first four sessions of a psychosocial program for smoking cessation.

Studies investigating pain or symptom distress used IVR as a distraction tool. Participants played an interactive game or experienced an interactive VE during the course of a painful or distressing medical intervention. Scenarios included shooting snowballs at targets whilst gliding through an icy canyon (Hoffman et al., 2008). All studies investigated a one-off treatment using IVR which was self-administered following brief training. The treatment using IVR and a control condition were delivered in the same or consecutive medical treatment sessions.

Across the studies the frequency of treatment using IVR varied between one-off to weekly sessions. The total time in IVR ranged from approximately 3-495 min (*Mdn* = 100 min). None of the studies considered the effect of presence on the outcome of treatment using IVR at follow-up.

3.1.2.4 IVR technology characteristics

Different VEs allowed different levels of participant interaction. This ranged from directly acting on the environment (e.g., purchasing a flight ticket; Rus-Calafell et al., 2013) to observing the environment (Price & Anderson, 2007). Various methods were used to facilitate interaction from body tracking devices (Meyerbröker et al., 2011) to a gamepad (Girard et al., 2009). Eight studies stated that projection was stereographic, two indicated that it was not (Malbos et al., 2013; Schneider & Hood, 2007). All studies reporting the display type used a head-mounted display with the exception of Chan et al. (2007) who used I-glasses. Meyerbröker et al. (2011) and Moldovan and David (2014) also used a computer automated virtual environment system.

3.1.2.5 Measures

Across the studies the most frequently used presence measure was the Presence Questionnaire (Witmer & Singer, 1998). Three studies used a modified version of this measure (Chan et al., 2007; Girard et al., 2009; Price & Anderson, 2007). Two studies (Meyerbröker et al., 2011; Price et al., 2011) used the Igroup Presence Questionnaire (Schubert, Friedmann, & Regenbrecht, 2001). Rus-Calafell et al. (2013) used the Presence and Reality Judgement Questionnaire (Baños et al., 2000). For the above questionnaires a total presence score is generated, with higher scores indicating a greater sense of presence. Riva et al. (2008) used the ITC-Sense of Presence Inventory (Lessiter, Freeman, Keogh, & Davidoff, 2001) whose four subscales each provide a separate score. Hoffman et al. (2000) and Hoffman et al. (2008) used a single item measure derived from the work of Hendrix and Barfield (1995).

Girard et al. (2009) and most studies investigating anxiety difficulties used domain-specific treatment outcome measures. Riva et al. (2008) only used domain-general measures. Malbos et al. (2013) and Moldovan and David (2014) included both types of measures. Studies investigating pain or symptom-distress used domain-general measures.

Table 2

Characteristics of Included Studies

Study	Clinical difficulty	Demographics [a]	Sample size	IVR treatment type	IVR treatment description	IVR frequency (F) and duration (D)	Technology	Measures	
								Presence	Treatment outcome
Malbos et al. (2013)	Anxiety: panic disorder with agoraphobia	**Age** = 44 (14) **Female** = 67% **Caucasian** = 100%	**Total** = 18 **IVR** = 18	Exposure	Experience nine VEs including a bridge and supermarket for a set duration. Investigator manipulates anxiogenic cues.	**F** = weekly **D** = 9 x 50-60 min	**Display** = HMD **Stereoscopy** = no **Tracker** = yes **Interaction** = act on the VE	PQ	ACQ ASI DASS MIA PA/W
Meyerbr-oker et al. (2011)	Anxiety: panic disorder with agoraphobia	**Age** = NS **Female** = NS **Caucasian** = NS	**Total** = 17 **IVR** = 11	Exposure	NS	**F** = NS **D** = 6 x NS	**Display** = HMD or CAVE **Stereoscopy** = yes **Tracker** = yes **Interaction** = NS	IPQ	ACQ MIA PDSS
Moldovan and David (2014)	Anxiety: social phobia or acrophobia or aviophobia	**Age** = NS **Female** = 47% **Caucasian** = NS	**Total** = 32 **IVR** = 16	Exposure	VEs NS. Administered according to an individualised fear hierarchy.	**F** = once **D** = 1 x 60 min	**Display** = HMD or CAVE **Stereoscopy** = NS **Tracker** = NS **Interaction** = NS	PQ	ABS II AQ BFNE FAM FAS LSAS SSPS STAI-Y
Price et al. (2011)	Anxiety: social phobia	**Age** = 40 (NS) [b] **Female** = 60% **Caucasian** = 54%	**Total** = 41[c] **IVR** = 41[c]	Exposure	VEs include a conference room, classroom and auditorium with an appropriate sized audience. Investigator manipulates audience reaction and questions according to treatment goals.	**F** = NS **D** = 4 x 20-30 min [b]	**Display** = HMD **Stereoscopy** = yes [b] **Tracker** = yes **Interaction** = act on the VE	IPQ	PRCS

(Table continues)

Study	Clinical difficulty	Demographics [a]	Sample size	IVR treatment type	IVR treatment description	IVR frequency (F) and duration (D)	Technology	Measures	
								Presence	Treatment outcome
Price and Anderson (2007)	Anxiety: specific phobia or panic disorder with flying as the main fear	**Age** = 39 (NS) **Female** = 85% **Caucasian** = 85%	**Total** = 36 **IVR** = 36	Exposure	Virtual plane with scenarios related to flying (e.g., sitting on the plane). Investigator controls progression through an individualised fear hierarchy allowing time for habituation.	**F** = 2 x a week **D** = 4 x 20-30 min [b]	**Display** = HMD **Stereoscopy** = yes [b] **Tracker** = NS **Interaction** = NS	PQ - adapted	FFI
Riva et al. (2008)	Anxiety: stress management for individuals with obesity	**Age** = NS **Female** = 100% **Caucasian** = NS	**Total** = 40 **IVR** = 15	Relaxation	Virtual island with relaxing experiences. Carry out relaxation exercises guided by a pre-recorded therapeutic narrative.	**F** = daily **D** = 2 x 60 min	**Display** = NS **Stereoscopy** = NS **Tracker** = NS **Interaction** = NS	ITC-SOPI	PANAS SAI VAS
Rus-Calafell et al. (2013)	Anxiety: aviophobia	**Age** = 37 (13) **Female** = 87% **Caucasian** = NS	**Total** = 15 **IVR** = 7	Exposure	Bedroom, airport and plane VEs involving tasks (e.g., pack a suitcase). Investigator manipulates conditions (e.g., turbulence). Follow a pre-established fear hierarchy, stay in each situation until habituation occurs.	**F** = 2 x a week **D** = 6 x 60-75 min	**Display** = HMD **Stereoscopy** = yes **Tracker** = yes **Interaction** = act on the VE	PRJQ	DEFAS FFS FFQ
Chan et al. (2007)	Pain	**Age** = 7 (2) **Female** = 13% **Caucasian** = NS	**Total** = 8 **IVR** = 8	Distraction	Play an interactive game: shoot ice-cream at a fox in an ice-cream factory.	**F** = once **D** = 1 x 15-20 min	**Display** = I-glasses **Stereoscopy** = yes **Tracker** = no **Interaction** = via a mouse	PQ - adapted	Faces scale

(Table continues)

Study	Clinical difficulty	Demographics [a]	Sample size	IVR treatment type	IVR treatment description	IVR frequency (F) and duration (D)	Technology	Measures	
								Presence	Treatment outcome
Hoffman et al. (2000)	Pain	**Age** = 28 (NS) **Female** = 8% **Caucasian** = NS	**Total** = 12 **IVR** = 12	Distraction	Explore a virtual kitchen and eat virtual candy (linked to real candy) and touch a virtual spider (linked to a toy spider).	**F** = once **D** = 1 x 3 min	**Display** = HMD **Stereoscopy** = yes **Tracker** = yes **Interaction** = act on the VE	1 question	5 questions rating pain
Hoffman et al. (2008)	Pain	**Age** = 27 (NS) **Female** = 0% **Caucasian** = NS	**Total** = 11 **IVR** = 11	Distraction	Play an interactive game: shoot snowballs at targets whilst gliding through an icy canyon.	**F** = once **D** = 1 x 3 min	**Display** = HMD **Stereoscopy** = yes **Tracker** = no **Interaction** = via a joystick	1 question	3 questions rating pain
Schneider and Hood (2007)	Symptom distress	**Age** = 54 (11) **Female** = 77% **Caucasian** = 91%	**Total** = 123 **IVR** = 107	Distraction	Play interactive games: deep sea diving; walk through an art museum; explore ancient worlds; solve a mystery. Change between games any time.	**F** = once **D** = 1 x 45-90 min	**Display** = HMD **Stereoscopy** = no **Tracker** = no **Interaction** = via a mouse	PQ	PFS SAI
Girard et al. (2009)	Tobacco addiction	**Age** = 44 (11) **Female** = 57% **Caucasian** = NS	**Total** = 91 **IVR** = 91	Therapeutic task	Find and crush virtual cigarettes (experimental condition) or virtual balls (control) located in a virtual castle.	**F** = weekly **D** = 4 x 30 min	**Display** = HMD **Stereoscopy** = yes **Tracker** = yes **Interaction** = act on the VE	PQ - adapted	FND

Note.
Presence measures: IPQ = Igroup Presence Questionnaire; ITC-SOPI = Independent Television Commission Sense of Presence Inventory; PQ = Presence Questionnaire; PRJQ = Presence and Reality Judgement Questionnaire.
Treatment outcome measures: ABS II = Attitudes and Beliefs Scale II; ACQ = Agoraphobia Cognitions Questionnaire; AQ = Acrophobia Questionnaire; ASI = Anxiety Sensitivity Index; BFNE = Fear of Negative Evaluations Scale – Brief version; DASS = Depression Anxiety Stress Scales; DEFAS = Danger Expectations and Flying Anxiety Scale; FAM = Flight Anxiety Modality Questionnaire; FAS = Flight Anxiety Situation Questionnaire; FFI = Fear of Flying Inventory; FFS = Fear of Flying Scale; FFQ = Fear of Flying Questionnaire; FND = Fagerstrom test for Nicotine Dependence – revised; LSAS = Liebowitz Social Anxiety Scale; MIA = Mobility Inventory for Agoraphobia; PA/W = Panic Attacks per Week; PANAS = Positive And Negative Affect Scale; PDSS = Panic Disorder Severity Scale; PFS = Revised Piper Fatigue Scale;

PRCS = Personal Report of Confidence as a Speaker; SAI = State Anxiety Inventory; SSPS = The Self Statements Towards Public Speaking Scale; STAI-Y = State-Trait Anxiety Inventory-Y Form; VAS = Visual Analogue Scale (to assess emotional states).
Other: IVR = immersive virtual reality; VE = virtual environment; HMD = head-mounted display; CAVE = computer automated virtual environment system.

[a] Age = *M* (*SD*).
[b] Data obtained through personal communication with Dr P. Anderson and Dr M. Price.
[c] Thirty-one participants from the IVR arm of an RCT. Ten participants from a clinical trial of IVR using fMRI.
● = Study included in meta-analysis.

3.2 Assessment of Methodological Quality

Methodological quality was assessed in six domains: reporting, external validity, internal validity-bias, internal validity-confounding, power and funding. The findings are summarised in Table 3.

3.2.1 Reporting

Seven studies provided aims or hypotheses about the effect of presence on the outcome of treatment using IVR (Hoffman et al., 2000; Malbos et al., 2013; Moldovan & David, 2014; Price & Anderson, 2007; Price et al., 2011; Rus-Calafell et al., 2013; Schneider & Hood, 2007). Most studies reported inclusion/exclusion criteria. Three reported participants' age, sex and ethnicity (Malbos et al., 2013; Price & Anderson, 2007; Schneider & Hood, 2007). All stated participants' clinical difficulty but the duration and severity of the difficulty were only reported by Chan et al. (2007), Girard et al. (2009), Malbos et al. (2013) and Rus-Calafell et al. (2013).

Treatment using IVR was adequately described in all but two studies (Meyerbröker et al., 2011; Moldovan & David, 2014). However, reporting of IVR technology characteristics varied. Six studies (Chan et al., 2007; Girard et al., 2009; Hoffman et al., 2000; Hoffman et al., 2008; Malbos et al., 2013; Schneider & Hood, 2007) measured cyber-sickness, a potential confound and adverse event.

With one exception (Riva et al., 2008), all studies described their measures prior to their results section. The administration of measures was documented less consistently. Means and random variability for presence and treatment outcome measures were reported by six studies (Chan et al., 2007; Girard et al., 2009; Hoffman et al., 2000; Malbos et al., 2013; Price & Anderson, 2007; Price et al., 2011). Statistical tests and actual probability values were fully reported by five

studies (Malbos et al., 2013; Meyerbröker et al., 2011; Price & Anderson, 2007; Price et al., 2011; Rus-Calafell et al., 2013) and partially reported by three studies (Girard et al., 2009; Hoffman et al., 2008; Schneider & Hood, 2007).

3.2.2 External validity

Four studies using a community sample (Girard et al., 2009; Malbos et al., 2013; Price et al., 2011; Rus-Calafell et al., 2013) indicated that participants were representative of the population from which they were recruited. Two of these (Price et al., 2011; Rus-Calafell et al., 2013) stated the proportion of those asked who agreed to participate. The treatment using IVR was administered by therapists in four studies (Moldovan & David, 2014; Price & Anderson, 2007; Price et al., 2011; Rus-Calafell et al., 2013) and an investigator in one study (Malbos et al., 2013). The former may be more representative of real-world procedures. Some studies took place at an independent psychology clinic (Girard et al., 2009; Price & Anderson, 2007), others were at a University psychology clinic (Price et al., 2011; Rus-Calafell et al., 2013). Whilst the former seems representative of a typical setting, representativeness of the latter seems less clear.

Apart from Chan et al. (2007), studies investigating IVR distraction seemed to use participants who were representative of the source population. Two studies reported the proportion of those asked who agreed to participate (Hoffman et al., 2000; Schneider & Hood, 2007). Treatment using IVR was administered during standard medical treatments at medical centres. These facilities seem typical for the source population. One study (Schneider & Hood, 2007) reported that treatment using IVR was facilitated by a research nurse. The facilitator was not stated by the other studies.

Four studies used manualised treatments (Meyerbröker et al., 2011; Price & Anderson, 2007; Price et al., 2011; Rus-Calafell et al., 2013) and three mentioned using affordable or commercially available equipment (Girard et al., 2009; Malbos et al., 2013; Schneider & Hood, 2007). This is important for enabling accessibility and replication of the intervention. All assessment criteria were unfulfilled by Riva et al. (2008).

3.2.3 Internal validity – bias

With three exceptions (Hoffman et al., 2000; Hoffman et al., 2008; Riva et al., 2008), studies indicated their analysis plan prior to reporting their findings. Some studies reported statistical decisions. These included using intent-to-treat analysis (Girard et al., 2009) and non-parametric statistics (Rus-Calafell et al., 2013). Statistical tests seemed appropriate although some studies conducting multiple comparisons (such as Malbos et al., 2013) did not consider adjustment for Type 1 error. Girard et al. (2009), Malbos et al. (2013) and Meyerbröker et al. (2011) ran different IVR conditions. These studies found no difference between the conditions in socio-demographics or measures of presence and outcome of the treatment using IVR. The studies combined the conditions to analyse the relationship between the latter two measures. This may have been due to the lack of statistical difference between the conditions on the above measures. However, none of the studies provided a rationale for this decision making it difficult to assess how appropriate it was.

Most studies carrying out a series of treatment sessions using IVR analysed a mean presence score across sessions (Girard et al., 2009; Malbos et al., 2013; Meyerbröker et al., 2011; Price et al., 2011). However, Price and Anderson (2007)

and Rus-Calafell et al. (2013) analysed the presence score from the first of a series of sessions. It is unknown whether this introduced bias as mixed results are reported for change in the level of presence experienced over time (Girard et al., 2009; Krijn, Emmelkamp, Olafsson, Schuemie, & van der Mast, 2007). Two studies were unclear whether a mean or one-off presence score was used for analysis (Moldovan & David, 2014; Riva et al., 2008).

Rus-Calafell et al. (2013) stated that treatment sessions were recorded to ensure fidelity of the intervention. Chan et al. (2007) reported qualitative data which suggests variation in treatment delivery but did not comment on this. Across the studies the scarcity of information regarding treatment compliance may lead to bias in interpreting findings. Most studies used valid measures. Riva et al. (2008) named well-established measures but did not describe or reference them. The single item presence measure used by Hoffman et al. (2000) and Hoffman et al. (2008) may be an incomplete assessment of the sense of presence since this sensation is commonly assumed to arise from several different elements of the IVR experience (Triberti et al., 2014).

3.2.4 Internal validity – confounding

Price et al. (2011) combined results from the IVR arm of an RCT and a before-after study which used identical treatment protocols. Apart from Price et al. (2011) studies which combined intervention groups for analysis of the relationship between presence and the outcome of treatment using IVR (Girard et al., 2009; Malbos et al., 2013; Meyerbröker et al., 2011) recruited participants for each group from the same population over the same period of time. Five studies investigating anxiety difficulties (Meyerbröker et al., 2011; Moldovan & David, 2014; Price &

Anderson, 2007; Riva et al., 2008; Rus-Calafell et al., 2013) randomly assigned participants to the IVR arm of the investigation. In the RCT arm of Price et al.'s (2011) study (see Anderson et al., 2013) and in two other studies (Moldovan & David, 2014; Rus-Calafell et al., 2013) randomisation was concealed until recruitment was complete. Studies which investigated pain or symptom distress used a within-subjects cross-over design. The order in which participants received treatment using IVR or the control condition was randomised.

Most studies investigating confounds found them to be negligible or statistically non-significant. Some studies failed to consider potential confounds. For example, some of Chan et al.'s (2007) participants (mean age 7 years) received parental support with completing measures. Despite findings that children experience presence differently to adults (Triberti et al., 2014), Hoffman et al. (2008) did not investigate differences between adult and child participants. Girard et al. (2009) combined IVR conditions for analysis, however, their results showed greater cyber-sickness in the experimental condition which could have skewed these results.

Three studies considered the effects of drop-out (Girard et al., 2009; Price & Anderson, 2007; Schneider & Hood, 2007). In other studies, attrition was either too small to effect the findings or did not occur.

3.2.5 Power

Only Schneider and Hood (2007) considered power. Their power calculation showed that their sample size was sufficient to test the effectiveness of the IVR intervention. Many of the other studies discussed the limitations of their small sample size.

3.2.6 Funding

One study (Girard et al., 2009) appeared to have a high risk of bias due to their funding source. Apart from Moldovan and David (2014) and Price et al. (2011) who did not report the source of funding, the other studies were funded by organisations which did not appear to have a vested interest in the results and therefore seemed at low risk of funding bias.

(Page left blank intentionally due to formatting)

Table 3

Assessment of Methodological Quality

Study	Reporting[a]	External validity[b]	Internal validity-bias[c]	Internal validity-confounding[d]	Power	Funding
Chan et al. (2007)	■	■	■	■	●	▲
Girard et al. (2009)	■	■	■	■	●	▲
Hoffman et al. (2000)	■	■	■	■	●	▲
Hoffman et al. (2008)	■	■	■	■	●	▲
Malbos et al. (2013)	■	■	■	■	●	▲
Meyerbroker et al. (2011)	■	■	■	■	●	▲
Moldovan and David (2014)	■	■	■	■	●	▲
Price and Anderson (2007)	■	■	■	■	●	▲
Price et al. (2011)	■	■	■	■	●	▲
Riva et al. (2008)	■	■	■	■	●	▲
Rus-Calafell et al. (2013)	■	■	■	■	●	▲
Schneider and Hood (2007)	■	■	■	■	●	▲

Note.
■ = more than two thirds of criteria met for this domain; ■ = more than one third of criteria met for this domain; ■ = less than one third of criteria met for this domain. Criteria were scored as 1 = criteria met, 0 = criteria unmet.
● = analysis performed; ● = no analysis performed.
▲ = low risk of bias; ▲ = unclear risk of bias; ▲ = high risk of bias.

[a] Reporting: aims and/or hypotheses, inclusion/exclusion criteria, demographics, clinical presentation, confounds/adverse events, measures, administration of measures, IVR intervention described, IVR frequency and duration, technology characteristics, means and random variability for measures, statistical test, actual probability value.

[b] External validity: representativeness of those asked to participate and the number prepared to participate, administration and setting, affordable/available technology.

[c] Internal validity – bias: data dredging, appropriate statistical tests, compliance with the intervention, accuracy of outcome measures.

[d] Internal validity – confounding: different groups recruited from the same population at the same time (where applicable), randomisation, allocation concealment, adjustment for confounds in the analysis, attrition accounted for.

3.3 Data Synthesis

3.3.1 The relationship between presence and the outcome of treatment using IVR

Seven studies provided data for meta-analysis (n = 311) of the overall effect size for the relationship between presence and the outcome of treatment using IVR. The analysis showed a small effect size with a weighted mean correlation of r = .28, p < .001, 95% CI [0.15, 0.39]. Figure 2 presents a Forest Plot of the analysis.

Unlike bivariate correlations, semi-partial correlations control for the effects of other variables in the model (Aloe & Becker, 2011). Therefore, combining bivariate and semi-partial correlations may confound results. Given this, a summary effect was computed for studies using bivariate correlations only (n = 184). This showed a small effect size with a weighted mean correlation of r = .27, p < .001, 95% CI [0.12, 0.40]. Figure 3 presents a Forest Plot of the analysis. Both analyses showed that a greater sense of presence during treatment using IVR is associated with better treatment response.

Statistical tests indicated that heterogeneity was unimportant for the first analysis (fixed effects: Q = 6.92, p = .33, I^2 = 13.34) and the second analysis (fixed effects: Q = 2.15, p = .71, I^2 = 0.00). Whilst statistical tests for heterogeneity are useful, the choice of model for the meta-analysis should primarily be based on whether studies share a common effect size (Borenstein et al., 2009). As the two analyses used a small number of studies which were not functionally equivalent a common effect size could not be assumed. Therefore, the use of a random-effects model was still considered appropriate.

The results reported by studies which were ineligible for meta-analysis are presented in Table 4. Riva et al. (2008) found that greater presence during treatment using IVR was associated with better treatment response. Hoffman et al. (2008)

found that participants whose presence rating was above the sample mean showed significant reductions in more aspects of pain following treatment using IVR than participants whose presence rating was below the mean. This was taken as indicating that greater presence during the treatment using IVR was associated with better treatment outcome. Hoffman et al. (2000) reported a trend towards such a relationship but it did not reach statistical significance. The authors attributed the statistical result to the small sample size. Chan et al. (2007) and Moldovan and David (2014) found no such relationship.

Table 4

Findings for the Relationship Between Presence and the Outcome of Treatment Using IVR for Studies Ineligible for Meta-Analysis

Source	**Analysis**	**Result**
Chan et al. (2007)	Correlation	○
Hoffman et al. (2000)	Correlation	○
Hoffman et al. (2008)	T-tests	●
Moldovan and David (2014)	Unclear	○
Riva et al. (2008)	Correlation	●

Note.
● = positive relationship between presence and the outcome of treatment using IVR.
● = positive relationship indicated between presence and the outcome of treatment using IVR.
○ = no relationship between presence and the outcome of treatment using IVR.

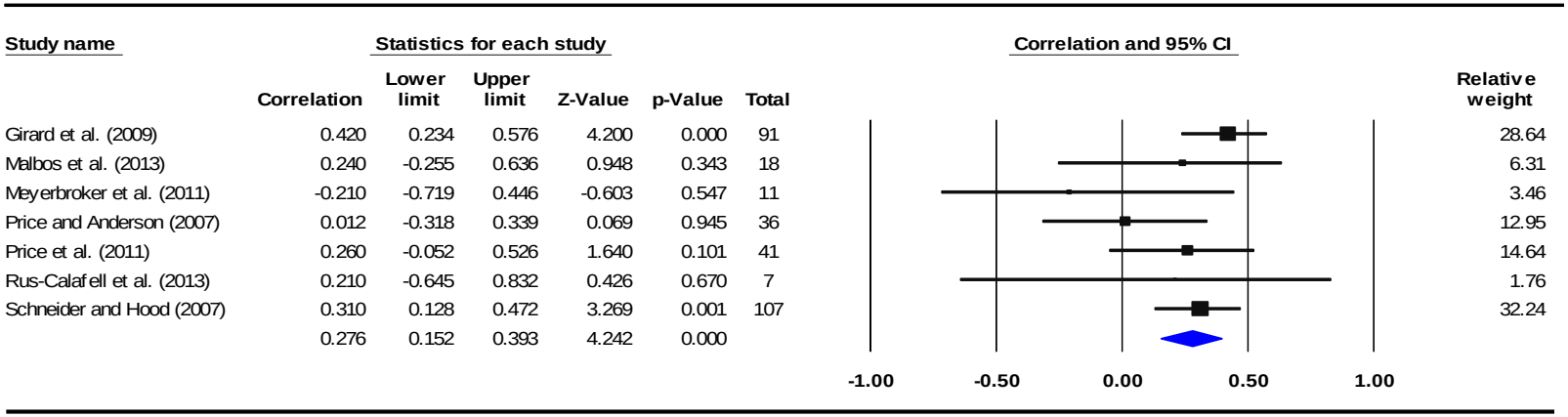


Figure 2. Forest plot of the relationship between presence and the outcome of treatment using IVR.

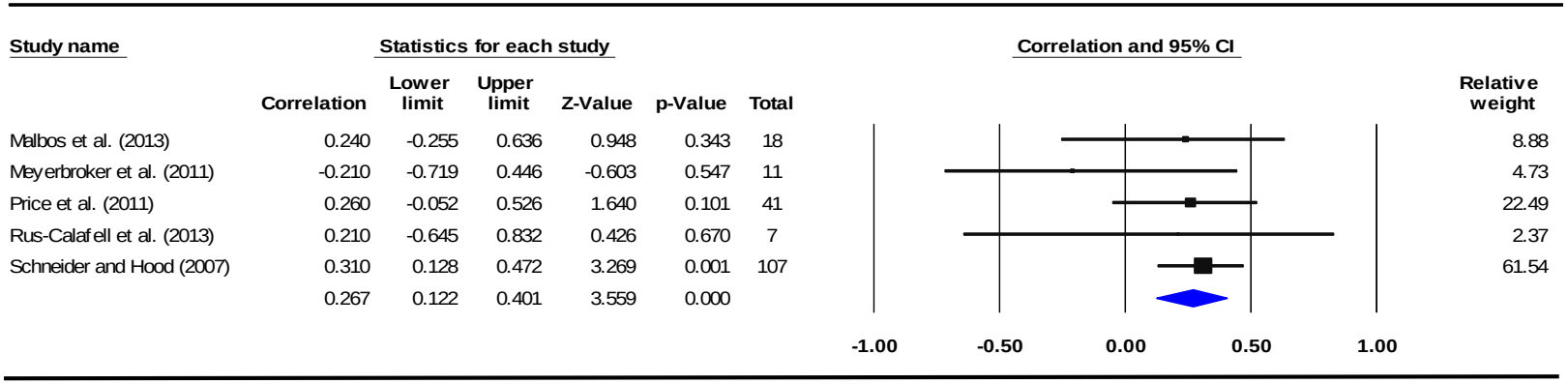


Figure 3. Forest plot of the relationship between presence and the outcome of treatment using IVR for studies reporting bivariate correlations.

3.3.2 Variables which potentially moderate the relationship between presence and the outcome of treatment using IVR

The following variables which may moderate the relationship between presence and the outcome of treatment using IVR were examined: participant characteristics, clinical difficulty, measures, type of treatment using IVR, IVR technology and sample size. The data presented concerns information reported by the studies or supplied by their corresponding authors. Below, where studies are reported to have found a relationship between presence and the outcome of treatment using IVR this refers to a positive association between these variables. No relationship refers to the finding that there was no relationship between these variables.

3.3.2.1 Participant characteristics

The studies were grouped according to whether or not they found a relationship between presence and the outcome of treatment using IVR. Participants' mean age was 42 years in studies which indicated a relationship and 33 years in studies indicating no relationship. The mean percentage of female participants was 59% for the former group of studies and 52% for the latter group of studies. Of the studies indicating a relationship only Schneider and Hood (2007) reported the percentage of Caucasian participants, which was 91%. Studies indicating no relationship existed used a mean of 80% Caucasian participants. Figure 4 presents the available data.

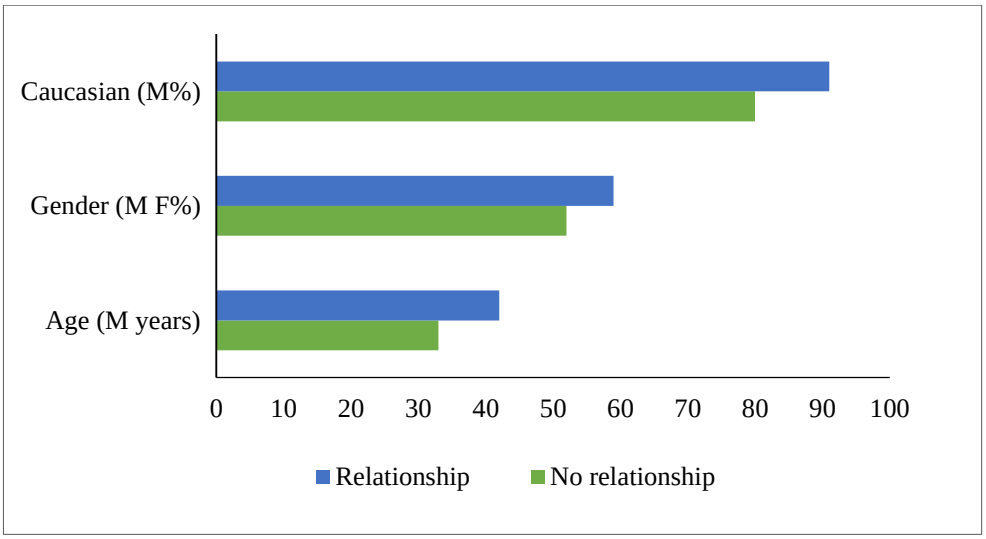


Figure 4. Participant characteristics grouped by finding for the relationship between presence and the outcome of treatment using IVR.

3.3.2.2 Clinical difficulty

One of the seven studies (Riva et al., 2008) which investigated participants with anxiety difficulties found a relationship between presence and the outcome of treatment using IVR. Studies investigating symptom distress or addiction also found that this relationship existed. Of the studies investigating pain, Hoffman et al.'s (2008) results indicated this relationship existed, Hoffman et al. (2000) found a trend towards it and Chan et al. (2007) found no such relationship.

3.3.2.3 Measures

The study using the ITC-Sense of Presence Inventory (Riva et al., 2008) and two of the six studies using the Presence Questionnaire (Girard et al., 2009; Schneider & Hood, 2007) reported a relationship between presence and the outcome of treatment using IVR. Hoffman et al. (2008), whose results indicated such a relationship, used a single item presence measure. Given this variation the type of presence measure does not appear to have a moderating effect. Three of the four

studies which found or indicated that this relationship existed (Hoffman et al., 2008; Riva et al., 2008; Schneider & Hood, 2007), and Hoffman et al. (2000) who reported a trend towards it, assessed the outcome of treatment using IVR with domain-general distress measures.

3.3.2.4 Treatment using IVR

Two of the four studies using IVR distraction (Hoffman et al., 2008; Schneider & Hood, 2007) indicated a relationship between presence and the outcome of treatment using IVR. Studies which used IVR to expose participants to relaxing scenarios or for therapeutic tasks to treat tobacco addiction also reported this outcome (Girard et al., 2009; Riva et al., 2008). The six studies using VRET (Malbos et al., 2013; Meyerbröker et al., 2011; Moldovan & David, 2014; Price & Anderson, 2007; Price et al., 2011; Rus-Calafell et al., 2013) and the other two studies using IVR distraction (Chan et al., 2007; Hoffman et al., 2000) did not find that this relationship existed.

Overall, descriptive statistics for the dose-response relationship suggest that fewer IVR treatment sessions with shorter total time in IVR is more likely to generate a relationship between presence and the outcome of treatment using IVR. Available data are presented in Figure 5.

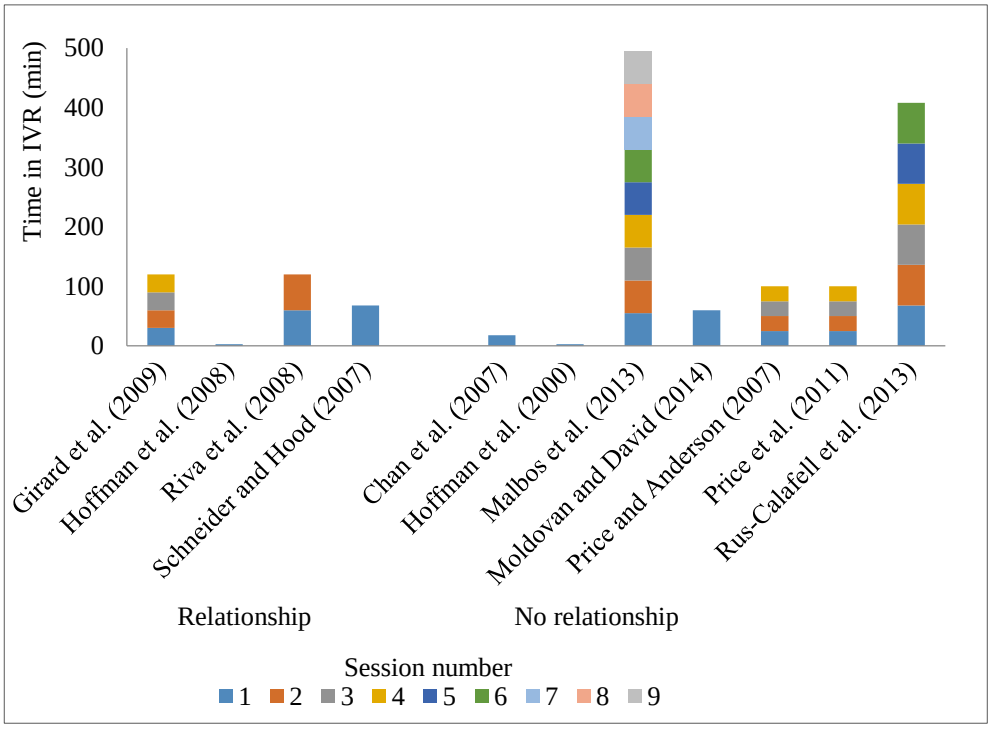


Figure 5. Dose-response relationship grouped by finding for the relationship between presence and the outcome of treatment using IVR.

3.3.2.5 Technology characteristics

Stereoscopy, tracking and type of interaction with the VE did not appear to be moderating factors. These technology characteristics differed across the studies.

3.3.2.6 Sample size

Sample size for studies which indicated a relationship between presence and the outcome of treatment using IVR ranged from 11-107 (*Mdn* = 53). Sample size for studies reporting no such relationship ranged from 7-41 (*Mdn* = 14).

4: Discussion

This review examined 12 studies which investigated the relationship between presence and the outcome of treatment using IVR. Data concerning this relationship were synthesised using meta-analysis and narrative synthesis. The discussion will first summarise the results. Then it will appraise the evidence used to generate the results and biases in the review process. Finally, implications for practice and research will be discussed.

4.1 Results of the Meta-Analyses and Narrative Synthesis

4.1.1 The relationship between presence and the outcome of treatment using IVR

Meta-analysis using seven studies (n = 311) found a small effect size for a relationship between presence and the outcome of treatment using IVR, r = .28, 95% CI [0.15, 0.39]. This showed that a greater level of presence during treatment using IVR was associated with better treatment response. As research in this field is in its infancy and limited the inclusion criteria for the review enabled all available data to be investigated. However, combining different statistical tests can introduce confounds (Borenstein et al., 2009). So to check the finding was robust a second analysis was carried out using only studies reporting bivariate correlations (n = 184). This also showed a small effect size, r = .27, 95% CI [0.12, 0.40] and the result supports the initial finding. Two (Hoffman et al., 2008; Riva et al., 2008) of the five studies ineligible for meta-analysis also indicated that this relationship existed.

Presence is commonly assumed to be a key mechanism underlying IVR treatment efficacy (Spagnolli et al., 2014). The meta-analyses support this. However, it is important to note these are correlational meta-analyses and therefore

the relationship between presence and the outcome of treatment using IVR does not necessarily imply that these variables are causally related. Only two studies included in the analysis found such a relationship (Girard et al., 2009; Schneider & Hood, 2007). These studies received a weighting of 28.64 and 32.24 respectively in the initial analysis and the latter received a 61.54 weighting in the second analysis. The differences in the characteristics of the included studies and their findings supports the rationale for examining variables which potentially moderate this relationship.

4.1.2 Variables which potentially moderate the relationship between presence and the outcome of treatment using IVR

4.1.2.1 Clinical difficulty and IVR treatment type

Riva et al. (2008) was the only study to find a relationship between presence and the outcome of treatment using IVR whose participants suffered from anxiety. However, unlike the other studies investigating anxiety, participants in Riva et al.'s (2008) study did not meet DSM-IV criteria for an anxiety disorder and were treated for stress using IVR relaxation. The other studies indicating a relationship treated pain, symptom distress and tobacco addiction using IVR for distraction or to administer a therapeutic task.

This finding could partly relate to participants' attentional focus during the treatment. Attention is required in order to experience a sense of presence (Witmer & Singer, 1998). Additionally, sustained attention during treatment is associated with improved treatment outcomes, whilst distraction during treatment is related to poorer treatment outcomes (Telch et al., 2004). It is possible that the more the individual directs their attention towards the VE, the less attentional capacity they

have available to process nociceptive signals or outside information (Cabas-Hoyos, Gutierrez-Martinez, Gutierrez-Maldonado, & Loreto, 2010; Hoffman et al., 2008).

IVR distraction is based on the premise that IVR exerts a powerful demand on attentional resources yet Hoffman et al. (2000) and Chan et al. (2007) did not find that presence was related to the outcome of treatment using IVR. However, the former study reported a trend towards this relationship. Chan et al. (2007) was the only study using just child participants. This may have affected their finding as children experience presence differently to adults (Triberti et al., 2014) and participants received parental support with filling out measures. Interestingly, nursing staff's observations suggested that participants who appeared less present in IVR became distracted and anxious when medical treatment began.

VRET treatments were interrupted by the administrator which might have reduced participants' attention towards the VE. The lack of relationship in VRET studies may also relate to Foa and Kozak's (1986) theory. This proposes that a phobic fear structure must be activated for emotional processing to occur. The importance of presence to elicit anxiety is supported by the finding that these factors are associated during the initial VRET session (Ling et al., 2014). However, once the fear structure is activated, controlled, prolonged, repeated exposure is required for habituation and extinction to occur. In support of this, research shows that greater time in VRET results in a more positive treatment outcome (Opriş et al., 2012; Powers & Emmelkamp, 2008). Therefore, presence may be necessary but insufficient by itself to lead to superior treatment outcome (Meyerbröker et al., 2011; Price & Anderson, 2007).

4.1.2.2 Outcome measures

The majority of studies reporting a relationship only used domain-general measures. Therefore, presence in IVR may improve general psychological distress rather than targeting disorder-specific symptoms. However, data did not allow comparison of domain-specific and domain-general measures within studies to consider this finding independently of other between-study variations.

Different measures of presence did not seem to moderate the relationship. However, these measures are based on different constructs of presence, may lack sensitivity to the specificities of therapy using IVR (Spagnolli et al., 2014) and single-item measures may prevent a complete evaluation of the sense of presence (Triberti et al., 2014). Moreover, the ability of self-report questionnaires to capture the subjective experience of presence has been questioned (Slater & Steed, 2000; Slater, 2004). Therefore, this finding must be interpreted with caution.

4.1.2.3 Dose-response relationship

Studies indicating a relationship between presence and the outcome of treatment using IVR generally used shorter and fewer IVR treatment sessions. However, Girard et al.'s (2009) study and an initial study of multiple sessions of IVR distraction treatment (Hoffman et al., 2001) found improved treatment outcome and increased presence with repeated immersion. Therefore, other factors may account for the finding of the current review. These could include fun (Hoffman et al., 2008), enjoyment, successful task completion (Girard et al., 2009) and a pleasant emotional experience during IVR. It is possible that such factors may sustain and enhance attention, generating presence and leading to improved treatment outcome in certain treatments using IVR such as IVR distraction and relaxation. For VRET, the

theoretical rationale and research findings support a null finding for the relationship between presence and the outcome of treatment using IVR in relation to IVR dosage (Opriş et al., 2012; Price & Anderson, 2007).

4.1.2.4 Other moderating variables

Compared to studies which did not find a relationship, studies which indicated a relationship used participants with a higher mean age and had a higher percentage of Caucasian and female participants. Research in this field is limited with mixed findings (Ling, Nefs, Brinkman, Qu, & Heynderickx, 2013) making the implications of this data difficult to determine.

The median sample size was 53 for studies which indicated a relationship and 14 for studies finding no relationship. This may suggest that studies finding no relationship were under-powered. Whilst literature discussed above makes this seem unlikely for VRET studies, it could help explain Chan et al.'s (2007) and Hoffman et al.'s (2000) findings. It seems less likely that sample size impacts on the relationship between presence and the outcome of treatment using IVR as sample size varied from 11-107 in studies indicating a relationship.

Contrary to Ling et al. (2014), technology characteristics did not appear to moderate the relationship. This finding may be affected by the small number of studies in the review and the paucity of data reported in these studies.

4.2 Overall Completeness and Applicability of Evidence

Across the 12 studies sample size varied from 7-107. The relatively small sample size of the majority of studies (in some cases below 10 participants) raises concern that they may have been under-powered to detect a true clinical difference.

Only 7 of the 12 studies reported statistical data for the relationship between presence and the outcome of treatment using IVR.

The majority of studies used validated measures. Combining categories of treatment outcome measures where both were included in one study limited interpretation of the findings. However, conducting separate analyses for each category was not feasible given the very small number of studies in each category. A further concern is that self-report measures are prone to demand characteristics. Objective (i.e., behaviour and physiological) treatment outcome measures were excluded from the current review as they are seldom used. It is interesting to note that Malbos et al. (2013) used objective measures and found that greater presence was associated with a more positive treatment outcome.

The meta-analysis addressed a broad question. This was considered appropriate given that research in this field is in its infancy. However, the included studies were highly heterogeneous. Factors including the use of other therapeutic components alongside the IVR treatment, paucity of data regarding treatment compliance and possible confounds from combining different conditions for analysis complicates interpretation. Investigation of some potentially moderating variables supports interpretation of the statistical summary by examining differences in effects between studies. But lack of data such as technology characteristics hinders completeness and applicability. Overall, external validity for sampling and the setting of IVR distraction treatment seemed greater than for VRET.

4.3 Quality of the Evidence

Random-effects meta-analysis was carried out with a small number of studies. Accordingly, random errors cannot be excluded. Furthermore, the estimate

of between-studies variance upon which the summary effect is partly based may be erroneous. Therefore the point estimate and confidence intervals may provide a false sense of assurance (Borenstein et al., 2009; Viechtbauer, 2005). However, it can be argued that a statistical summary may still be superior to an ad hoc summary with unknown properties (Borenstein et al., 2009). It was assumed that Hoffman et al.'s (2008) results indicated a relationship. This was not based on direct statistical data and therefore may have led to biased conclusions.

More than one third of reporting criteria were fulfilled by all except two studies. Apart from considerations of power, external validity received the lowest quality rating overall with four studies meeting less than one third of criteria for this domain. Internal validity-bias received the highest quality rating overall with seven studies meeting more than two thirds of criteria. Whilst more than two thirds of criteria for internal validity-confounding were met by six studies, three met less than one third. Only one study performed a power analysis. One study appeared at risk of bias in relation to funding. It is likely that not all assessment items are of equal weight and importance. However, based on the assessment carried out, the methodological quality of the studies appeared highly variable.

4.4 Potential Biases in the Review Process

Established guidelines (Higgins & Green, 2011) and assessment tools (Downs & Black, 1998) were consulted throughout the review process. The review was limited to peer-reviewed, English language publications which may have excluded valuable contributions from case studies and grey literature. Study selection and methodological quality assessment was performed by one reviewer.

Although efforts were made to contact corresponding authors to obtain missing data only authors from one study provided information.

4.5 Implications for Practice

Until recently, treatment using IVR has been criticised for being inaccessible and expensive (Glantz, Rizzo, & Graap, 2003). But as technology advances, its application in clinical settings becomes more feasible. Findings from the present review are tentative. The data is limited, heterogeneous and of varied quality meaning generalisation should be carried out with caution. Although preliminary, findings suggest that sense of presence may be efficacious for reducing pain and symptom distress using IVR distraction. A similar outcome was found for relaxation using IVR to treat stress and IVR therapeutic tasks to treat tobacco addiction. It has been hypothesised that factors including fun and inducing a positive emotional experience may enhance and sustain attention, leading to a greater sense of presence and improved treatment outcome for the above treatments.

Findings for VRET do not indicate that increasing presence alone will benefit treatment efficacy. However, presence may be important in initial stages of VRET (Ling et al., 2014). Individualising the VE, reducing outside noise and distraction by taking practical steps such as avoiding interruptions from the investigator may help generate the sense of presence required.

4.6 Implications for Research

Research should continue to work towards an explicit definition of presence at a conceptual level. Current differences in the operationalisation of presence in questionnaires hinders the quality, robustness and replicability of findings.

This review tentatively indicates that presence may play a different role in different types of treatment using IVR as a conduit for treatment effectiveness. This fits with previous research (Price & Anderson, 2007; Triberti et al., 2014). Large scale, adequately powered, controlled studies whose primary aim is to examine this relationship are required across different types of treatment using IVR to increase the robustness of the evidence-base. This will require separating treatment using IVR from other therapeutic components to eliminate their contribution to the relationship. The current investigation could be extended using behavioural avoidance tests (to assess generalisability of treatment using IVR to real life phobic scenarios) and examining factors which may sustain and enhance presence.

Greater understanding of variables which potentially moderate presence and treatment using IVR is required. The current review suggests that such research should include further examination of demographic, measurement and technology factors. Moreover, the mechanisms and components of effective treatment using IVR need elucidating. The role of attention and its links with presence have emerged as key questions to address.

Given that individuals with psychological difficulties and health care providers may prefer treatment using IVR to alternative options (Gorini & Riva, 2008; Wiederhold & Wiederhold, 2007) developing the evidence-base to make treatment outcomes better predictable is crucial. Clarity and comprehensiveness in research process and reporting will strengthen this body of literature. Without this, advancement and adoption of treatment using IVR will be limited given the high standard of research for more established interventions (McCann et al., 2014).

4.7 Conclusion

Findings from the review suggest that a greater sense of presence during treatment using IVR is associated with a more positive treatment outcome for IVR distraction treatment, relaxation treatment and treatment for tobacco addiction. By itself, presence does not seem to be sufficient to affect the outcome of VRET. These conclusions are tentative in light of the heterogeneity and variation in methodological quality of the included studies as well as limitations of the review process. Large scale, adequately powered, controlled studies are needed for robust conclusions along with further examination of moderating factors.

References

Aloe, A. M. (2014). An empirical investigation of partial effect sizes in meta-analysis of correlational data. *The Journal of General Psychology*, *141*(1), 47–64. doi:10.1080/00221309.2013.853021

Aloe, A. M., & Becker, B. J. (2011). An effect size for regression predictors in meta-analysis. *Journal of Educational and Behavioral Statistics*, *37*(2), 278–297. doi:10.3102/1076998610396901

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. *Journal of Consulting and Clinical Psychology*, *81*(5), 751–760. doi:10.1037/a0033559

Baños, R. M., Botella, C., Garcia-Palacios, A., Villa, H., Perpiña, C., & Alcañiz, M. (2000). Presence and reality judgment in virtual environments: A unitary construct? *CyberPsychology & Behavior*, *3*(3), 327–335.

Bordnick, P. S., Carter, B. L., & Traylor, A. C. (2011). What virtual reality research in addictions can tell us about the future of obesity assessment and treatment. *Journal of Diabetes Science and Technology*, *5*(2), 265–271. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3125915&tool=pmcentrez&rendertype=abstract

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: John Wiley & Sons.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2014). Comprehensive Meta Analysis (Version 3.2) [computer software]. Englewood NJ, USA: Biostat, Inc. Retrieved from www.meta-analysis.com

Bowman, D. A., & McMahan, R. P. (2007). Virtual reality: How much immersion is enough? *Computer*, *40*(7), 36–43.

Cabas-Hoyos, K., Gutierrez-Martinez, O., Gutierrez-Maldonado, J., & Loreto, D. (2010). The illusion of presence influences VR distraction: Effects on cold-pressor pain. *Annual Review of CyberTherapy and Telemedicine*, *8*, 123–126.

Chan, E. A., Chung, J. W., Wong, T. K., Lien, A. S., & Yang, J. Y. (2007). Application of a virtual reality prototype for pain relief of pediatric burn in Taiwan. *Journal of Clinical Nursing*, *16*(4), 786–793. doi:10.1111/j.1365-2702.2006.01719.x

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (Vol. 2). Hillsdale, N.J. : L. Erlbaum Associates. doi:10.1234/12345678

Corey, D. M., Dunlap, W. P., & Burke, M. J. (1998). Averaging correlations: Expected values and bias in combined pearson rs and fisher's z transformations. *The Journal of General Psychology*, *125*(3), 245–261. doi:10.1080/00221309809595548

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology & Community Health*, *52*(6), 377–384. doi:10.1136/jech.52.6.377

Emmelkamp, P. M., Bruynzeel, M., Drost, L., & van der Mast, C. A. G. (2001). Virtual reality treatment in acrophobia: A comparison with exposure in vivo. *CyberPsychology & Behavior*, *4*(3), 335–339. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11710257

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

Gerardi, M., Cukor, J., Difede, J., Rizzo, A., & Rothbaum, B. O. (2010). Virtual reality exposure therapy for post-traumatic stress disorder and other anxiety disorders. *Current Psychiatry Reports*, *12*(4), 298–305. doi:10.1007/s11920-010-0128-4

Girard, B., Turcotte, V., Bouchard, S., & Girard, B. (2009). Crushing virtual cigarettes reduces tobacco addiction and treatment discontinuation. *CyberPsychology & Behavior*, *12*(5), 477–483. doi:10.1089/cpb.2009.0118

Glantz, K., Rizzo, A., & Graap, K. (2003). Virtual reality for psychotherapy: Current reality and future possibilities. *Psychotherapy: Theory, Research, Practice, Training*, *40*(1), 55–67. Retrieved from http://psycnet.apa.org/journals/pst/40/1-2/55/\nhttp://032912b.membershipsoftware.org/libdocuments/Virtual_Reality_Uses.pdf

Gold, J. I., Kim, S. H., Kant, A. J., Joseph, M. H., & Rizzo, A. S. (2006). Effectiveness of virtual reality for pediatric pain distraction during IV placement. *CyberPsychology & Behavior*, *9*(2), 207–212. doi:10.1089/cpb.2006.9.207

Gorini, A., & Riva, G. (2008). Virtual reality in anxiety disorders: The past and the future. *Expert Review of Neurotherapeutics*, *8*(2), 215–33. doi:10.1586/14737175.8.2.215

Gregg, L., & Tarrier, N. (2007). Virtual reality in mental health: A review of the literature. *Social Psychiatry and Psychiatric Epidemiology*, *42*, 343–354. doi:10.1007/s00127-007-0173-4

Grewe, P., Kohsik, A., Flentge, D., Dyck, E., Botsch, M., Winter, Y., … Piefke, M. (2013). Learning real-life cognitive abilities in a novel 360°-virtual reality supermarket: A neuropsychological study of healthy participants and patients

with epilepsy. *Journal of Neuroengineering and Rehabilitation*, *10*(1), 42–57. doi:10.1186/1743-0003-10-42

Hendrix, C., & Barfield, W. (1995). Presence in virtual environments as a function of visual and auditory cues. In *Proceedings Virtual Reality Annual International Symposium '95* (pp. 74–82).

Higgins, J. P. T., & Green, S. (Eds.). (2011). *Cochrane handbook for systematic reviews of interventions Version 5.1.0 [updated March 2011]*. The Cochrane Collaboration. Retrieved from www.cochrane-handbook.org

Hodges, L. F., Kooper, R., Meyer, T. C., Rothbaum, B. O., Opdyke, D., de Graaff, J. J., … North, M. M. (1995). Virtual environments for treating the fear of heights. *Computer*, *28*(7), 27–34.

Hoffman, H. G., Garcia-Palacios, A., Carlin, A., Furness, T. A., & Botella-Arbona, C. (2003). Interfaces that heal: Coupling real and virtual objects to treat spider phobia. *International Journal of Human-Computer Interaction*, *16*(2), 283–300.

Hoffman, H. G., Patterson, D. R., & Carrougher, G. J. (2000). Use of virtual reality for adjunctive treatment of adult burn pain during physical therapy: A controlled study. *The Clinical Journal of Pain*, *16*(3), 244–250.

Hoffman, H. G., Patterson, D. R., Carrougher, G. J., Nakamura, D., Moore, M., Garcia-Palacios, A., & Furness, T. A. (2001). The effectiveness of virtual reality pain control with multiple treatments of longer durations: A case study. *International Journal of Human-Computer Interaction*, *13*(1), 1–12. doi:10.1207/S15327590IJHC1301_1

Hoffman, H. G., Patterson, D. R., Seibel, E., Soltani, M., Jewett-Leahy, L., & Sharar, S. R. (2008). Virtual reality pain control during burn wound debridement in the

hydrotank. *The Clinical Journal of Pain*, *24*(4), 299–304. doi:10.1097/AJP.0b013e318164d2cc

Hoffman, H. G., Sharar, S. R., Coda, B., Everett, J. J., Ciol, M., Richards, T., & Patterson, D. R. (2004). Manipulating presence influences the magnitude of virtual reality analgesia. *Pain*, *111*, 162–168. doi:10.1016/j.pain.2004.06.013

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. *Behaviour Research and Therapy*, *42*(2), 229–239. doi:10.1016/S0005-7967(03)00139-6

Krijn, M., Emmelkamp, P. M. G., Olafsson, R. P., Schuemie, M. J., & van der Mast, C. (2007). Do self-statements enhance the effectiveness of virtual reality exposure therapy? A comparative evaluation in acrophobia. *CyberPsychology & Behavior*, *10*(3), 362–370. doi:10.1089/cpb.2006.9943

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-sense of presence inventory. *Presence*, *10*(3), 282–297. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6790390

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., … Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration. *BMJ*, *339*. doi:10.1136/bmj.b2700

Ling, Y., Nefs, H. T., Brinkman, W.-P., Qu, C., & Heynderickx, I. (2013). The relationship between individual characteristics and experienced presence. *Computers in Human Behavior*, *29*(4), 1519–1530. doi:10.1016/j.chb.2012.12.010

Ling, Y., Nefs, H. T., Morina, N., Heynderickx, I., & Brinkman, W. P. (2014). A meta-analysis on the relationship between self-reported presence and anxiety in virtual reality exposure therapy for anxiety disorders. *PloS ONE*, *9*(5). doi:10.1371/journal.pone.0096144

Mahrer, N. E., & Gold, J. I. (2009). The use of virtual reality for pain control: A review. *Current Pain and Headache Reports*, *13*, 100–109. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19272275

Malbos, E., Rapee, R. M., & Kavakli, M. (2013). A controlled study of agoraphobia and the independent effect of virtual reality exposure therapy. *The Australian and New Zealand Journal of Psychiatry*, *47*(2), 160–168. doi:10.1177/0004867412453626

Malloy, K. M., & Milling, L. S. (2010). The effectiveness of virtual reality distraction for pain reduction: A systematic review. *Clinical Psychology Review*, *30*(8), 1011–1018. doi:10.1016/j.cpr.2010.07.001

McCann, R. A., Armstrong, C. M., Skopp, N. A., Edwards-Stewart, A., Smolenski, D. J., June, J. D., … Reger, G. M. (2014). Virtual reality exposure therapy for the treatment of anxiety disorders: An evaluation of research quality. *Journal of Anxiety Disorders*, *28*(6), 625–631. doi:10.1016/j.janxdis.2014.05.010

Melzack, R., & Wall, P. D. (1965). Pain mechanisms: A new theory. *Science*, *150*, 971–979. doi:10.1126/science.150.3699.971

Meyerbröker, K., & Emmelkamp, P. M. G. (2010). Virtual reality exposure therapy in anxiety disorders: A systematic review of process-and-outcome studies. *Depression and Anxiety*, *27*(10), 933–944. doi:10.1002/da.20734

Meyerbröker, K., Morina, N., Kerkhof, G., & Emmelkamp, P. M. G. (2011). Virtual reality exposure treatment of agoraphobia: A comparison of computer automatic

virtual environment and head-mounted display. *Annual Review of CyberTherapy and Telemedicine*, *9*(1), 41–45.

Moldovan, R., & David, D. (2014). One session treatmet of cognitive and behavioral therapy and virtual reality for social and specific phobias. Preliminary results from a randomized clinical trial. *Journal of Cognitive and Behavioral Psychotherapies*, *14*(1), 67–83.

Moon, J., & Lee, J. H. (2009). Cue exposure treatment in a virtual environment to reduce nicotine craving: A functional MRI study. *CyberPsychology & Behavior*, *12*(1), 43–45. doi:10.1089/cpb.2008.0032

Morris, L. D., Louw, Q. A., & Grimmer-Somers, K. (2009). The effectiveness of virtual reality on reducing pain and anxiety in burn injury patients: A systematic review. *The Clinical Journal of Pain*, *25*(9), 815–826. doi:10.1097/AJP.0b013e3181aaa909

Opriş, D., Pintea, S., García-Palacios, A., Botella, C., Szamosközi, Ş., & David, D. (2012). Virtual reality exposure therapy in anxiety disorders: A quantitative meta-analysis. *Depression and Anxiety*, *29*(2), 85–93. doi:10.1002/da.20910

Park, C.-B., Choi, J.-S., Park, S. M., Lee, J.-Y., Jung, H. Y., Seol, J.-M., … Kwon, J. S. (2014). Comparison of the effectiveness of virtual cue exposure therapy and cognitive behavioral therapy for nicotine dependence. *Cyberpsychology, Behavior and Social Networking*, *17*(4), 262–267. doi:10.1089/cyber.2013.0253

Parsons, T. D., & Rizzo, A. A. (2008). Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry*, *39*(3), 250–261. doi:10.1016/j.jbtep.2007.07.007

Powers, M. B., & Emmelkamp, P. M. G. (2008). Virtual reality exposure therapy for anxiety disorders: A meta-analysis. *Journal of Anxiety Disorders*, *22*(3), 561–569. doi:10.1016/j.janxdis.2007.04.006

Price, M., & Anderson, P. L. (2007). The role of presence in virtual reality exposure therapy. *Journal of Anxiety Disorders*, *21*(5), 742–751. doi:10.1016/j.janxdis.2006.11.002

Price, M., Mehta, N., Tone, E. B., & Anderson, P. L. (2011). Does engagement with exposure yield better outcomes? Components of presence as a predictor of treatment response for virtual reality exposure therapy for social phobia. *Journal of Anxiety Disorders*, *25*(6), 763–770. doi:10.1016/j.janxdis.2011.03.004

Riva, G. (2005). Virtual reality in psychotherapy: Review. *CyberPsychology & Behavior*, *8*(3), 220–230. doi:10.1089/cpb.2005.8.220

Riva, G., Manzoni, M., Villani, D., Gaggioli, A., & Molinari, E. (2008). Why you really eat? Virtual reality in the treatment of obese emotional eaters. *Studies in Health Technology and Informatics*, *132*, 417–419.

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. *CyberPsychology & Behavior*, *6*(5), 467–476. doi:10.1089/109493103769710497

Rus-Calafell, M., Gutiérrez-Maldonado, J., Botella, C., & Baños, R. M. (2013). Virtual reality exposure and imaginal exposure in the treatment of fear of flying: A pilot study. *Behavior Modification*, *37*(4), 568–590. doi:10.1177/0145445513482969

Sanchez-Vives, M. V., & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience*, *6*(4), 332–339. doi:10.1038/nrn1651

Schneider, S. M., & Hood, L. E. (2007). Virtual reality: A distraction intervention for chemotherapy. *Oncology Nursing Forum*, *34*(1), 39–46.

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence*, *10*(3), 266–281.

Schuemie, M. J., van der Straaten, P., Krijn, M., & van der Mast, C. A. (2001). Research on presence in virtual reality: A survey. *CyberPsychology & Behavior*, *4*(2), 183–201. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11710246

Sharar, S. R., Carrougher, G. J., Nakamura, D., Hoffman, H. G., Blough, D. K., & Patterson, D. R. (2007). Factors influencing the efficacy of virtual reality distraction analgesia during postburn physical therapy: Preliminary results from 3 ongoing studies. *Archives of Physical Medicine and Rehabilitation*, *88*(12), 43–49. doi:10.1016/j.apmr.2007.09.004

Slater, M. (2004). How colorful was your day? Why questionnaires cannot assess presence in virtual environments. *Presence*, *13*(4), 484–493.

Slater, M., & Steed, A. (2000). A virtual presence counter. *Presence*, *9*(5), 413–434. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6788093

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence*, *6*(6), 603–616. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=9716&login.asp&site=ehost-live&scope=site

Spagnolli, A., Bracken, C. C., & Orso, V. (2014). The role played by the concept of presence in validating the efficacy of a cybertherapy treatment: A literature review. *Virtual Reality, 18*, 13–36. doi:10.1007/s10055-013-0241-x

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., … Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ, 342*. doi:10.1136/bmj.d4002

Sun, X., Briel, M., Walter, S. D., & Guyatt, G. H. (2010). Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. *BMJ, 340*. doi:http://dx.doi.org/10.1136/bmj.c117

Telch, M. J., Valentiner, D. P., Ilai, D., Young, P. R., Powers, M. B., & Smits, J. A. J. (2004). Fear activation and distraction during the emotional processing of claustrophobic fear. *Journal of Behavior Therapy and Experimental Psychiatry*, *35*(3), 219–232.

Triberti, S., Repetto, C., & Riva, G. (2014). Psychological factors influencing the effectiveness of virtual reality-based analgesia: A systematic review. *Cyberpsychology, Behavior and Social Networking*, *17*(6), 335–345. doi:10.1089/cyber.2014.0054

Viechtbauer, W. (2005). Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics*, *30*(3), 261–293. Retrieved from http://jeb.sagepub.com/content/30/3/261.short

Villani, D., Riva, F., & Riva, G. (2007). New technologies for relaxation: The role of presence. *International Journal of Stress Management*, *14*(3), 260–274. doi:10.1037/1072-5245.14.3.260

Villani, D., & Riva, G. (2008). Presence and relaxation: A preliminary controlled study. *PsychNology Journal*, *6*(1), 7–25.

Vincelli, F. (1999). From imagination to virtual reality: The future of clinical psychology. *CyberPsychology and Behavior*, *2*(3), 241–248. Retrieved from http://online.liebertpub.com/doi/abs/10.1089/109493199316366

Wiederhold, B. K., Davis, R., & Wiederhold, M. D. (1998). The effects of immersiveness on physiology. *Studies in Health Technology and Informatics*, *58*, 52–60. doi:10.3233/978-1-60750-902-8-52

Wiederhold, M. D., & Wiederhold, B. K. (2007). Virtual reality and interactive simulation for pain distraction. *Pain Medicine*, *8*, 182–188. doi:10.1111/j.1526-4637.2007.00381.x

Wismeijer, A. A. J., & Vingerhoets, A. J. J. M. (2005). The use of virtual reality and audiovisual eyeglass systems as adjunct analgesic techniques: A review of the literature. *Annals of Behavioral Medicine*, *30*(3), 268–278. doi:10.1207/s15324796abm3003_11

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence*, *7*(3), 225–240.

Part 2: Empirical Study

A randomised trial exploring the relationship between visuo-spatial perspective-taking, avatar embodiment and the ability to cultivate compassion using virtual reality and mental imagery

Abstract

Aims

This study explored whether allocentric visuo-spatial perspective-taking (VSPT) ability or sense of embodiment affect the ability to cultivate self-compassion in self-critical individuals using an immersive virtual reality (IVR) or mental imagery (MI) intervention. Change in self-compassion and self-criticism following the intervention was examined. Participants' experience of the intervention and effects related to practicing imagining the intervention for two weeks were investigated.

Method

This was a parallel-groups, stratified randomisation, non-blinded study. Healthy adults high in trait self-criticism were randomly assigned to a one-off IVR (n = 20) or analogue MI (n = 20) intervention. Participants completed an allocentric VSPT task pre-intervention, an embodiment measure post-intervention and a state self-compassion and self-criticism measure pre-intervention, post-intervention and at two week follow-up. Ease of recall, frequency of practice and image vividness ratings were also completed at follow-up.

Results

Allocentric VSPT ability and embodiment were unrelated to change in state self-compassion or self-criticism following the interventions. State self-criticism reduced after both interventions but state self-compassion increased only after MI. The IVR intervention was experienced more negatively. Ease of recalling the MI intervention was positively related to allocentric VSPT ability. Reduction in state self-criticism after the IVR intervention was associated with greater image vividness.

Conclusions

Rather that VSPT ability or embodiment it may be the experience of the intervention that influenced state self-compassion and self-criticism. The MI intervention was more efficacious in cultivating self-compassion however the efficacy of the IVR intervention may be developed by addressing aspects that were experienced negatively. This study would benefit from replication and extension to investigate variations of the interventions.

1: Introduction

1.1 Perspective-Taking

The ability to mentally adopt another person's perspective is an important aspect of human social cognition. Two lines of research have emerged in this area. One has investigated visuo-spatial perspective-taking (VSPT), the ability to see the world from another person's perspective, taking into account what they see and how they see it (Surtees, Apperly, & Samson, 2013). In egocentric VSPT representations are made relative to the viewer (Blanke, 2012). Allocentric VSPT involves representations relative to another person or object (Thakkar & Park, 2010). The other line of research has focused on social cognition which is concerned with Theory of Mind (ToM; Premack & Woodruff, 1978) or mentalising (Frith & Frith, 2006), the ability to represent and attribute mental states to others. Drawing on the VSPT literature, Frith and De Vignemont (2005) propose that visuo-spatial and social processes could be unified, suggesting that the adoption of either an egocentric or allocentric stance supports ToM and mentalising. In the former the individual represents the other's mental states in relation to themselves, in the latter the other's mental states are represented independently from the self.

A growing body of research has linked social perspective-taking and VSPT (Clements-Stephens, Vasiljevic, Murray, & Shelton, 2013; Hamilton, Kessler, & Creem-Regehr, 2014; Kessler & Wang, 2012). Hamilton, Brindley, and Frith (2009) found that children's ability to complete an allocentric ToM task predicted performance on allocentric VSPT tasks. Deficits on both these tasks have been found in individuals with Autistic Spectrum Disorder (Pearson, Ropar, & Hamilton, 2013) and may be found in those with an egocentric VSPT deficit (Frith & De Vignemont, 2005). In schizophrenia impairments have been found in both

allocentric VSPT and cognitive empathy (Langdon, Coltheart, Ward, & Catts, 2001; Thakkar & Park, 2010; Thirioux, Tandonnet, Jaafari, & Berthoz, 2014). Here, cognitive empathy, the capacity to understand the experience and associated mental state of others (Eisenberg, Wentzel, & Harris, 1998), is closely akin to allocentric ToM. Furthermore, neuroimaging studies have identified brain areas commonly engaged in both social perspective-taking and VSPT (Frith & Frith, 2006; Lambrey, Doeller, Berthoz, & Burgess, 2012; Schurz, Aichhorn, Martin, & Perner, 2013; Spreng, Mar, & Kim, 2009).

1.2 Self-Compassion and Perspective-Taking

Self-compassion is defined as self-kindness/warmth and a sense of common humanity (Gilbert, 2009a; Neff, 2003). This form of self-to-self relating is negatively associated with psychopathology (Barnard & Curry, 2011; MacBeth & Gumley, 2012; Neff, Rude, & Kirkpatrick, 2007). Social mentality theory proposes that soothing and reassuring (compassionate) responses from significant others stimulate specific pathways in the brain which generate a sense of being safe, loved and lovable. This builds emotional memories of the self in relationships with others which are copied for self-to-self relating (Gilbert, 2000, 2005). When individuals self-regulate with self-compassion they elicit soothing neuro-affective responses of well-being, safeness and social connectedness, similar to those which might be stimulated by a supportive other (Gilbert, 2005, 2009b). Thus, self-compassion involves taking the stance of a compassionate other towards ourselves (Germer & Neff, 2013).

Given this, it seems that allocentric social perspective-taking may be an important component of self-compassion (Gilbert & Irons, 2005). The individual is

required to adopt another person's compassionate viewpoint and understand or perceive the situation the way they do (Thirioux, Mercier, Blanke, & Berthoz, 2014; Tomasello, Kruger, & Ratner, 1993). In support of this, Neff and Pommier (2013) found that self-compassion was positively associated with social perspective-taking ability (which they defined as "stepping into another's shoes" so that one understands their point of view, p.162) in healthy samples of undergraduates, adults and practicing meditators. It therefore seems possible that the ability to take the stance of a compassionate other is linked to VSPT ability.

1.3 Self-Compassionate Mental Imagery

Mental imagery (MI) has a powerful capacity to impact on emotions, physiology and cognition (Hackmann & Holmes, 2004) and can be used for therapeutic benefit (Kaplan & Epstein, 2012; Zhang, Yu, & Barrett, 2014). MI exercises designed to cultivate self-compassion have been developed to treat maladaptive self-criticism. This form of negative self-judgement and evaluation is a pervasive feature of psychopathology (Gilbert & Irons, 2005) and predicts poor outcomes in psychotherapy (Rector, Bagby, Segal, Joffe, & Levitt, 2000). According to social mentality theory, individuals who self-regulate using self-criticism have a heightened threat-focused system (Gilbert, 2009b). Based on the conceptualisation of self-compassion, MI exercises designed to cultivate self-compassionate self-regulation involve imagining compassion flowing out from oneself to another, and from another into oneself (Gilbert & Choden, 2013; Gilbert & Irons, 2005). These MI exercises have been found to activate a soothing-affiliation system in the brain (Longe et al., 2010; Rockliff, Gilbert, McEwan, Lightman, & Glover, 2008). This

facilitates an internal compassionate relationship with the self in which self-warmth and self-soothing down-regulate the threat-focused system (Gilbert & Procter, 2006).

Therapeutic benefit from MI may be enhanced by practice (Kaplan, Epstein, & Sullivan-Smith, 2015; Neff & Pommier, 2013) and the ability to generate vivid mental images (Gaesser & Schacter, 2014; Kelly, Zuroff, Foa, & Gilbert, 2010). Additionally, being able to easily recall self-reassuring, compassionate images is likely to increase self-reassuring/compassionate self-to-self relating (Brewin, 2006; Gilbert, Baldwin, Irons, Baccus, & Palmer, 2006).

Most individuals report increased feelings of self-compassion following self-compassionate MI interventions (Gilbert & Irons, 2004; Gilbert & Procter, 2006; Kelly, Zuroff, & Shapira, 2009; Shahar et al., 2014). However, some highly self-critical individuals find attempts to develop images and feelings of self-compassion difficult or distressing (Gilbert & Irons, 2005; Gilbert & Procter, 2006; Rockliff et al., 2008). Such individuals may not have access to memories of being affectionately cared for and their self-care abilities may have been under-stimulated (Gilbert, 2014a). Consequently, they may find it easy to generate vivid self-critical images but struggle to access self-reassuring images (Gilbert et al., 2006). Furthermore, the ability to imagine compassion flowing from oneself to another, and from another into oneself requires flexible, allocentric spatial representation, computed by the hippocampus (Bird, Capponi, King, Doeller, & Burgess, 2010; Hartley et al., 2007). Individuals who are less able to use allocentric spatial processing may have an impaired ability to generate novel images, impacting on their capacity to carry out self-compassionate MI. Therefore, therapy to cultivate self-compassion may be advanced by finding alternative ways to access and experience this form of self-to-self relating.

1.4 Immersive Virtual Reality

Recently, immersive virtual reality (IVR) has been investigated as an alternative therapeutic tool to MI. This technology immerses the participant in a computer-generated, interactive virtual environment (Gregg & Tarrier, 2007). By conveying rich, realistic perceptual experiences it removes the individual's need to rely on internal mental images. To date, IVR and MI have mostly been compared for exposure treatments where they have been found to be similarly efficacious (Gamito et al., 2010; Rus-Calafell et al., 2013; Wallach, Safir, & Bar-Zvi, 2009). Investigation of IVR as a tool for cultivating compassion has just begun. Gillath, McCall, Shaver, and Blascovich (2008) found that self-reported compassion predicted compassionate behaviour in IVR. They argue that this suggests correspondence between compassionate behaviour in IVR and the real-world, indicating that IVR is efficacious for fostering such real-world behaviour. Further to this, Falconer et al. (2014) investigated a one-off IVR intervention designed to cultivate self-compassion in a healthy sample of females high in trait self-criticism. Favourable changes were observed in state measures of self-criticism and to a lesser extent, self-compassion. This suggests that IVR holds promise as an alternative to MI to cultivate self-compassion and counter self-criticism.

1.5 Embodiment

IVR technology enables an individual to inhabit an avatar and become immersed in their perspective. IVR participants can identify with avatars whose bodies have a different appearance, for example, different age or race (Banakou, Groten, & Slater, 2013; Peck, Seinfeld, Aglioti, & Slater, 2013; Won, Bailenson, Lee, & Lanier, 2015) and different perceptual experiences (Ahn, Le, & Bailenson,

2013) from their own. Therefore IVR is uniquely capable of facilitating social perspective-taking as participants can embody another person's perspective and experience (Raij, Kotranza, Lind, & Lok, 2009). This virtual experience has been found to impact on the participant's attitudes and behaviour (Ahn et al., 2013; Raij et al., 2009).

Virtual embodiment is obtained through the illusion that the virtual body is in fact the participant's body (Petkova & Ehrsson, 2008; Slater, Spanlang, Sanchez-Vives, & Blanke, 2010). The illusion requires a first person perspective and sensorimotor and visuomotor synchrony (i.e., the participant's field of view and body movements map onto the virtual body and are updated in real time; Kokkinara & Slater, 2014; Maselli & Slater, 2013). The latter can be enhanced by a virtual mirror in which the participant can view the reflection of their self-avatar (González-Franco, Pérez-Marcos, Spanlang, & Slater, 2010).

In Falconer et al.'s (2014) novel intervention a participant was firstly embodied as an adult avatar and was required to respond compassionately to a crying child avatar. Secondly the participant was re-embodied as the child avatar and experienced a recorded replay of their earlier interaction as the adult avatar delivering the compassionate response. By exploiting the embodiment of avatars in IVR and allocentric processing this intervention objectifies the notion of self-compassion as the participant gives compassion to themselves. Therefore, a participant's sense of embodiment and allocentric VSPT ability may relate to their ability to benefit from the intervention. Interestingly, Falconer et al. (2014) found that change in self-compassion was present regardless of the strength of the sense of embodiment as measured by the Virtual Reality Experience Questionnaire (VREQ; Falconer et al., 2014). This finding requires confirmation. The VREQ also includes

questions concerning other aspects of the participant's experience of the scenario (e.g., as the child, was the individual reassured by the adult?) which were not investigated by Falconer et al. (2014). It is possible that these aspects may also relate to the efficacy of the intervention, and furthermore could relate to the efficacy of a similar MI intervention.

1.6 The Current Study

Exploring factors which may affect the ability to cultivate self-compassion in self-critical individuals using IVR and MI interventions and examining the influence of these interventions on self-compassion and self-criticism would advance understanding of how to nurture self-compassionate self-to-self relating. To this end, this study investigated Falconer et al.'s (2014) IVR scenario and an analogue MI scenario as one-off interventions using a healthy sample of highly self-critical individuals. Allocentric VPST ability and sense of embodiment, factors which may affect the efficacy of the interventions, were explored. Change in state self-compassion and self-criticism following each intervention was examined. For comprehensive exploration of possible affecting factors, participants' experience of the interventions, and the effect of practicing imagining the scenario, image vividness and ease of recall over the two weeks post-intervention were also assessed.

The following hypotheses and exploratory questions derived from the literature were investigated:

- Hypothesis 1: Allocentric VSPT ability will be positively associated with change in state self-compassion and may affect state self-criticism. Difference between the conditions is uncertain.

- Question 1: Does sense of embodiment in IVR relate to change in state self-compassion or state self-criticism?
- Question 2: Does sense of embodiment in IVR relate to allocentric VSPT ability?
- Hypothesis 2: Following the intervention state self-compassion may increase and state self-criticism may be affected. Difference in the effect of the conditions is uncertain.
- Question 3: Are the interventions experienced differently?
- Question 4: Does frequency of practice, experiencing the image vividly or recalling it with ease for two weeks after the intervention relate to allocentric VSPT ability or sense of embodiment in IVR? Does this differ between the conditions?
- Hypothesis 3: Frequency of practice, experiencing the image vividly or recalling it with ease for two weeks after the intervention will lead to greater increase in state self-compassion and may affect state self-criticism. Difference in the effect of the conditions is uncertain.

2: Method

2.1 Design

This was a single centre, parallel-groups, stratified randomisation, non-blinded study conducted in the UK. Participants were randomly assigned to either an IVR or MI condition in a 1:1 ratio. Assignment was performed using a computer-generated random schedule in permuted blocks of two within gender strata (male and female). Gender balancing was used due to gender-bias in tendency to adopt an

egocentric verses allocentric visuo-spatial perspective (Gardner, Sorhus, Edmonds, & Potts, 2012; Mohr, Rowe, & Blanke, 2010).

2.2 Sample Size

Given that the study was exploratory and to balance feasibility, clinical and statistical considerations sample size was generated for a medium effect for the interaction between condition and state self-compassion and state self-criticism scores. Therefore, the power calculation was based on Hypothesis 2. This was considered an appropriate sample size calculation as change in the state measures was a common component of the hypotheses and exploratory questions. The power calculation was calculated using G*Power Version 3.1.7 (Faul, Erdfelder, Lang, & Buchner, 2007). The parameters were set to: statistical test = ANOVA: repeated-measures within-between interaction, effect size $f = .2$ (Cohen, 1992), $\alpha = .05$, $\beta = .8$, number of groups = 2 (between subjects: IVR, MI), number of measurements = 3 (within subjects: assessed pre-intervention, post-intervention and at two week follow-up [FU]), correlation among repeated measures = .5, non-sphericity correction = 1. The overall sample size required was 42.

2.3 Participants

A sample size of 40 participants (20 per condition) was achieved. Participants' demographic and baseline data are presented in Table 1. To participate individuals were required to be aged 18 years or over and score above 20 on the Inadequate Self subscale of the Forms of Self-Criticising/Attacking and Self-Reassuring Scale (FSCRS; Gilbert, Clarke, Hempel, Miles, & Irons, 2004). This was the upper third of scores in a large undergraduate sample in a previous pilot study at

the same urban university as the current study and was taken as indicating high trait self-criticism. Exclusion criteria were having ever received treatment for mental illness or brain damage. After recruitment began previous participation in an IVR study at the university was added to the exclusion criteria as Falconer et al.'s (2014) study was taking place at the same location. Individuals were excluded if this data was missing. One potential participant was excluded due to technical problems with the IVR.

2.4 Procedure

The study occurred between July and December 2014. It was part of a joint project with another Trainee Clinical Psychologist researcher (see Holden, n.d., for the other researcher's study and Appendix B for details of each researchers contribution to the joint project). Questionnaire measures were administered via online surveys constructed using Opinio, Version 6.8 (2014). As this did not allow counter-balancing administration was in the orders listed below.

2.4.1 Recruitment

Participants were recruited using self-selection though online advertisements (i.e., Facebook and the university electronic newsletter, group email system and psychology subject pool). Advertisements included a link to an eligibility survey. The survey questions concerned contact details, demographic and inclusion criteria. Opinio, Version 6.8 (2014) automatically generated a participant number when the survey was started. All information was stored in password protected spreadsheets against participant numbers which were marked according to the participant's eligibility. Contact details were stored separately to all other data. The author used

information in the contact details spreadsheet to email all participants. Emails to eligible participants provided further information about the study (Appendix C) and invited them to attend a one-off session at the university virtual reality lab. Emails to ineligible participants informed them that they had not been selected to continue. Simultaneously, using information in a data spreadsheet, the other researcher entered eligible participant numbers into the random assignment schedule. Attaining an equal sized sample for the conditions occurred spontaneously. Figure 1, prepared in accordance with CONSORT guidelines (Schulz, Altman, Moher, & Group, 2010), summarises the flow of participants through the study.

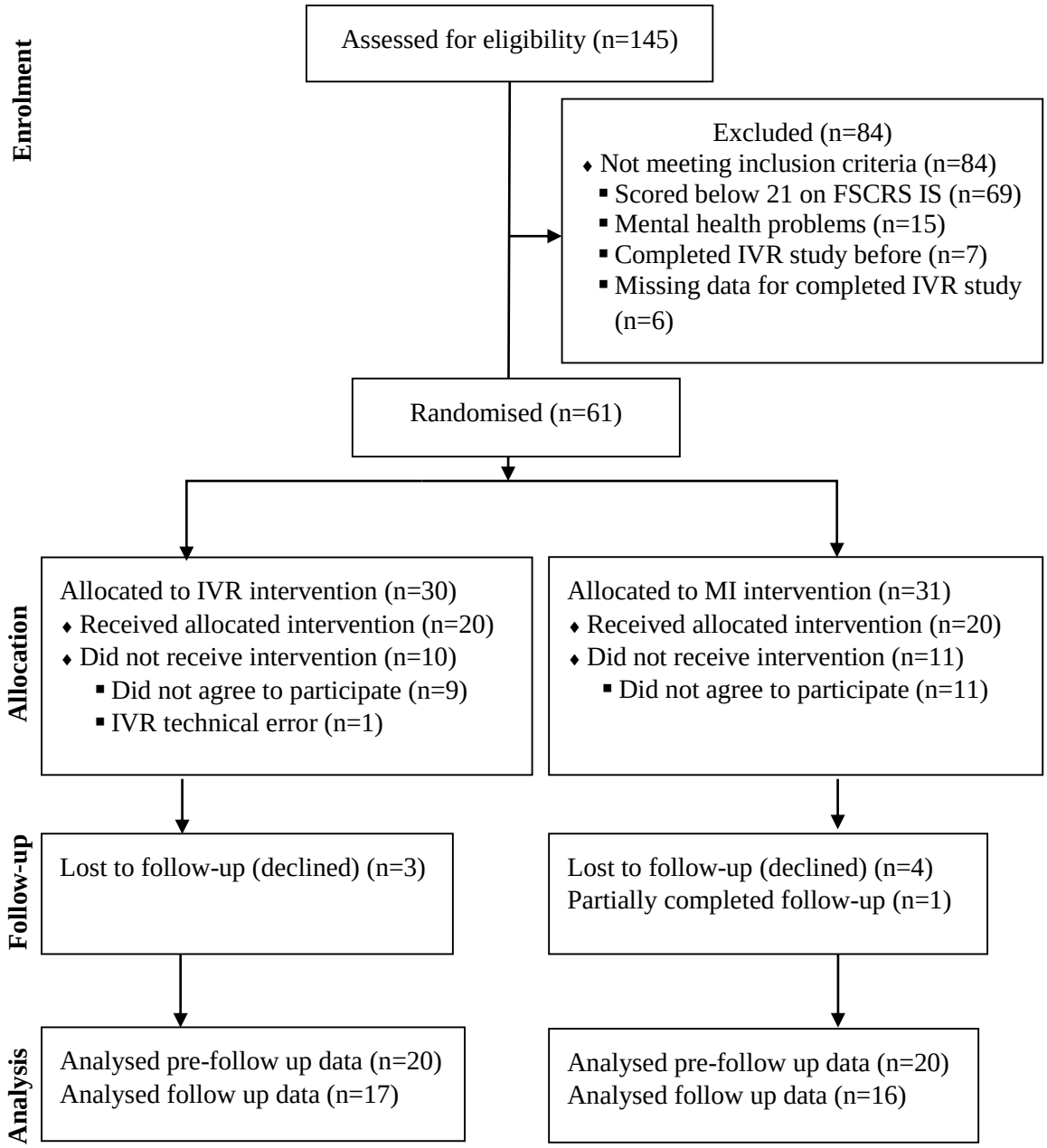


Figure 1. Flow of participants through the study.

2.4.2 Assessment

Pre-intervention participants completed, in order, the State Self-Compassion and Self-Criticism Scale questionnaire (SCCS; Falconer, King, & Brewin, 2015), an egocentric VSPT screening task, a topographical perception task and a topographical memory task. Immediately post-intervention participants completed an embodiment questionnaire and the SCCS. Two weeks post-intervention they were emailed a link

to a FU survey. This consisted of ease of recall, frequency of practice and image vividness questionnaires and the SCCS. Questionnaires were self-paced, tasks were timed. The eligibility and FU survey were completed at a time and place the participant chose. They completed all other assessments on a computer in the lab separated from the researchers by curtains which divided the room.

2.4.3 Intervention

Upon entering the lab the participant was informed of their assigned condition and asked to read an information sheet about the relevant intervention (Appendix C). After completing the pre-intervention assessments, they were given a sheet of paper listing the following three step compassionate response:

> Step 1 Acknowledge: *It's not nice when things happen to us that we don't like. It's really upset you hasn't it?*
> Step 2 Redirect the child's attention: *Sometimes when we are sad it's helpful to think of someone who loves us or is kind to us.*
> Step 3 Memory activation: *Can you think of someone who loves you or is kind to you? What might they say to you now that would make you feel better?*

The participant had five minutes alone to memorise the response as best they could to deliver at their own pace, in a slow, soft, compassionate manner during the intervention. They rehearsed once with a researcher. In the IVR condition participants put on a body-tracking suit and head mounted display and the IVR

system was calibrated. During both interventions the researchers and participant remained on separate set sides of the curtains.

2.4.3.1 IVR intervention

The virtual environment accurately replicated the participant's area of the lab: three blue walls, black curtains instead of a fourth wall, a brown door and sensors around the room. It additionally contained a full-length mirror and a stool. Participants saw themselves as an adult avatar of their gender. They heard a three minute audio recording which asked them to carry out exercises whilst looking in the mirror and then move freely for 30 seconds. This was intended to help them become accustomed to the virtual environment and embody the avatar. Then the visual scene faded out and a researcher reminded them of the compassionate response and the next stage of the intervention. The visual scene faded in and the participant saw a crying child avatar (of their gender) in the room. At their own pace and with their own accompanying movements the participant delivered the compassionate response. The child's disposition changed after each step of the response indicating a gradual improvement in mood: start, hunched over and crying into their hands; after step 1, hands lowered and sniffling; after step 2, less hunched over and stopped sniffling; after step 3, upright with their head elevated.

The visual scene faded out and a researcher reminded the participant what would happen next. The visual scene faded in and the participant saw themselves as the child avatar. They heard the same audio recording to facilitate embodiment of this avatar. The visual scene faded out. When it faded in the participant saw their adult avatar in the room. They experienced a recorded replay of themselves as the adult avatar delivering the compassionate response (including their body movements and voice). Participants were asked to look at and listen to their response from their new, child perspective. Participants typically spent 15 minutes in the IVR intervention and 90 minutes in the lab in total. See Figure 2 for images of the IVR intervention.

Figure 2. A female participant embodied in the adult avatar observing the child avatar's disposition during different stages of the compassionate response.

2.4.3.2 MI intervention

Participants sat in the location of the IVR intervention. They heard a pre-recorded script through headphones (see Appendix D for the script). This asked them to close their eyes and guided them to imagine the above scenario. The room, exercises and child were described. The following differed from the IVR intervention: participants imagined the characters were sat opposite them; the adult's

appearance was not described; as the adult they were prompted to imagine saying a step of the response and then informed of the child's reaction for each step of the response in turn; as the child they were informed what the adult had said and asked to imagine their adult-self saying it to them for each step of the response in turn. Participants had a keyboard on their lap. At regular intervals they were asked to press a key to continue. This enabled self-pacing. Participants typically spent 15 minutes in the MI intervention and 60 minutes in the lab in total.

2.4.4 Post-intervention

Participants completed the post-intervention measures. Then they were asked to practice imaging the scenario at least once a day for the following two weeks and for verbal consent to receive a text message reminder about this every other day. They were informed that after two weeks they would be emailed a FU survey and upon completing it they would be entered into a prize draw. Participants were thanked, debriefed and given a debrief sheet to take away (Appendix C).

2.4.5 IVR equipment

The virtual environment was created using Autodesk 3ds Max software and virtual avatars from Rocketbox studios. It was implemented with Unity 3D 4 game engine and an nVidia Quadro4000 graphics card. The head mounted display was nVisor SX111 with 1280 x 1024 pixel resolution and 102^0 horizontal field of view. Head tracking was via a 6-DOF Intersense IS-900. Body tracking was via Natural Point's Optitrack system using 12 V100 infrared Optitrack cameras to track 37 light reflective passive markers attached to a body suit.

2.5 Ethics

The study was approved by the Ethics Committee of the UCL Division of Psychology and Language Sciences as an amendment to an approved study (Appendix E). It was conducted according to the principles laid out in the Declaration of Helsinki. Informed consent was obtained prior to completing the eligibility survey and prior to the intervention (Appendix E).

2.6 Measures

2.6.1 Trait self-criticism

The 22-item, self-report FSCRS (Gilbert et al., 2004) asks participants to rate on a 5-point Likert scale (ranging from 0 = *not at all like me* to 4 = *extremely like me*) how they typically think and react when things go wrong for them. The scale has two subscales related to self-criticism, Inadequate Self (IS; nine items, e.g., "I think that I deserve my self-criticism") and Hated Self (HS; five items, e.g., "I stop caring about myself"), and a Reassuring Self subscale (RS; eight items, e.g., "I still like being me"). Internal reliability for each subscale is Cronbach's αs = .90, .86, and .86, respectively. The IS subscale was used in this study as HS items are more strongly endorsed in clinical populations (Longe et al., 2010).

2.6.2 Egocentric VSPT

The egocentric VSPT task, adapted from Ratcliff (1979) and presented using PsychoPy, Version 1.80.03 (Pierce, 2009, 2014), was used to screen for an egocentric VSPT deficit. Images of a manikin at 0^0 or 180^0 rotation, front or back facing with a black disk on one hand and white disk on the other hand are presented (Figure 3). After 2 practice trials, 16 images, 2 of each combination, are shown in a

pre-set randomised order. Participants press ‘f’ or ‘j’ on a keyboard to indicate if the black circle is on the manikin’s left or right hand respectively. Performance is the number of correct responses. Before and after the task there are three trials of an image of a black circle and white circle on either side of a vertical line. The same response keys are used to indicate where the black circle is. This checks understanding of left and right. Self-paced instructions are given before each set of trails. Each trial is shown for 30 seconds or until a response key is pressed. Non-response is scored as incorrect.

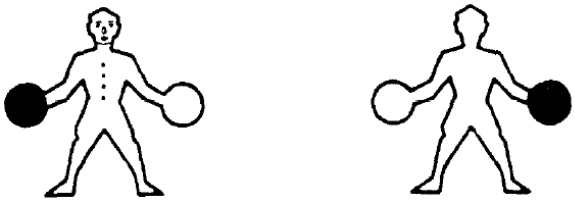


Figure 3. Example stimuli from the egocentric VSPT task.

2.6.3 Allocentric VSPT

Allocentric VSPT was assessed using the Four Mountains topographical perception and topographical memory tasks (Hartley et al., 2007). These were presented consecutively, in the stated order, using PsychoPy, Version 1.80.13 (Pierce, 2009, 2014). Both tasks present an image of four mountains in a landscape (target) and four foil images in a grid. Viewpoint and non-spatial features vary between foil and target images (Figure 4). Participants press the keyboard response key which corresponds to the grid positon of the foil image that matches the topography of the target: ‘e’ for top left, ‘x’ for bottom left, ‘i’ for top right, ‘m’ for bottom right. No image is repeated in a task. Self-paced instructions and three examples are given before the first task. Both tasks have 15 trials.

Topographical perception task: The target is presented at the top of the screen and the four foil images are presented in a grid below. Each trial of the task is shown for 30 seconds or until a response key is pressed.

Topographical memory task: The target image is presented in isolation for eight seconds. Then the screen turns blank for two seconds. Then the four foil images are presented for 30 seconds or until a response key is pressed.

Performance on each task is the number of correct responses. Non-response is scored as incorrect. Hartley et al. (2007) found that eight healthy individuals (Age M = 28.4, SD = 2.9) achieved a mean score of 14 on the topographical perception task and 13 on the topographical memory task.

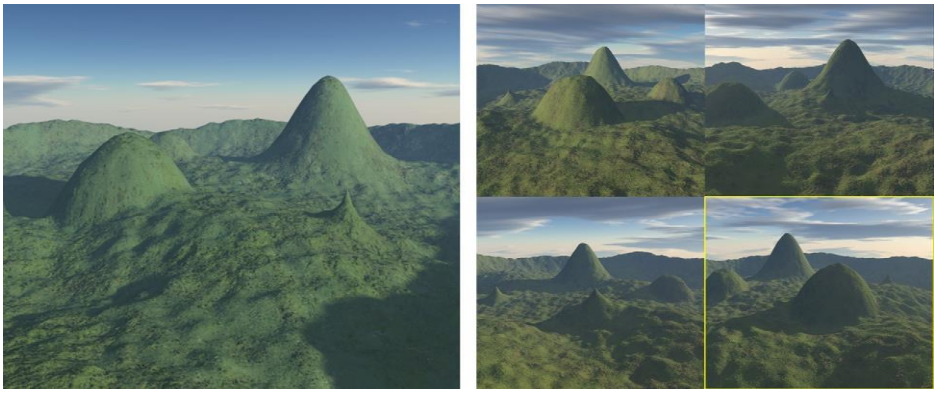


Figure 4. Example stimuli from the Four Mountains tasks.

2.6.4 Embodiment

Participants in the IVR condition completed the VREQ (Falconer et al., 2014). This consists of statements related to the sense of embodiment in IVR and other aspects of the participant's experience of the IVR scenario. All statements are rated on a 7-point Likert scale (ranging from -3 = *not at all* to +3 = *very much so*, or similar descriptions). The score for three statements which directly indicate the

illusion of body ownership (González-Franco et al., 2010; e.g., "As an adult I felt as if the body I saw when I looked in the mirror was my own body") and statements concerning aspects of the IVR scenario comparable with the MI scenario were assessed individually.

Participants in the MI condition completed the Mental Imagery Experience Questionnaire (MIEQ), an adaptation of the VREQ developed for this study. It consists of 16 modified VREQ statements relevant to the participant's experience of the MI scenario (e.g., "My talking to the child had a positive effect on their state"). The rating scale is as the VREQ. The score for each statement was assessed individually. Psychometric properties of these measures are not established.

2.6.5 State self-compassion and self-criticism

The SCCS (Falconer et al., 2015) consists of eight scenarios (e.g., "A third job rejection letter in a row arrives in the post"). Participants imagine each scenario is currently occurring and rate on a 7-point Likert scale (ranging from 1 = *not at all* to 7 = *highly*) the extent to which they would react to themselves in a harsh, critical, contemptuous, soothing, reassuring and compassionate manner. Across scenarios the former three ratings are summed to generate a state self-criticism score and the latter three are summed to generate a state self-compassion score. The authors report internal reliability for each score as Cronbach's αs = .87, and .91, respectively on a five scenario version of this measure.

2.6.6 Follow-up measures

To assess ease of recall participants rated on a 5-point Likert scale (ranging from 1 = *very difficult* to 5 = *very easy*) "How easy was it recall the scenario?" To

assess frequency of practice participants rated "Over the past two weeks, how often have you recalled the image generated by the scenario?" on a 7-point scale (ranging from 1 = *never* to 7 = *more than once a day*). These rating scales were developed for the study. Their psychometric properties are not established.

Image vividness was assessed using an adaptation of Kelly et al.'s (2010) Imagery Vividness assessment. Participants rated the extent to which they were able to bring to mind the following aspects of an image over a specified period of time: hearing the voice, seeing the facial expression, visualising the gestures, picturing the image, giving compassion and receiving compassion. Responses are rated on a 5-point scale (ranging from 1 = *perfectly clear and as vivid as in person* to 5 = *no image at all, you only 'know'*). A mean score across items is calculated. Kelly et al. (2010) found Cronbach's αs = .85, .90, and .92, respectively after each week of a three week self-compassionate MI intervention.

2.7 Data Analysis

Analysis was conducted using SPSS Statistics, Version 22. All data were checked for normality and transformed using a square root transformation if skew > 2, the Kolmogorov-Smirnov test was significant at $p < .01$, and the histogram appeared to deviate markedly from a normal distribution. For all analyses significance was set at $p < .05$, and adjusted for familywise error rate using Bonferroni correction (in accordance with Linacre, n.d.). Prior to analysis outliers on each variable were winsorised. Scores above the ninety-ninth percentile or below the first percentile were transformed to these percentiles respectively (Ghosh & Vogt, 2012). For correlational analysis the effect size of the correlation coefficient was interpreted as small $r = .1$, medium $r = .3$, and large $r = .5$ (Cohen, 1992).

Difference in demographic and pre-intervention data between the IVR and MI conditions, and between eligible individuals who did or did not complete the intervention was examined using *t* tests or chi-squared tests.

Descriptive data for the egocentric VSPT task was extracted to check for floor effects indicating a deficit in this ability. For each VSPT task Pearson's correlation was used to check whether the number of correct responses and time taken to complete the task were related. Allocentric VSPT was analysed using a measure of pure allocentric memory calculated by subtracting the topographical perception task score from the topographical memory task score. To evaluate the effect of VSPT ability Pearson's correlations were run between pure allocentric memory scores and both state self-compassion and state self-criticism change scores. Correlations were calculated for each condition separately then compared using Zou's method (Zou, 2007). This constructs approximate confidence intervals for the difference between two correlations (notation: 95% CI_z). If the interval includes zero there is no evidence of a statistically significant difference between them (Singer, 2013).

The VREQ and MIEQ were analysed non-parametrically due to their restricted scales. To assess the effect of embodiment in IVR Spearman's correlations were run between VREQ statements which indicate the illusion of body ownership (González-Franco et al., 2010) and pure allocentric memory scores, state self-compassion and state self-criticism change scores. To examine participants' experience of the intervention difference between corresponding statements on the VREQ and MIEQ was assessed using Mann-Whitney tests. Effect size was calculated as *r* using the formula, $r = Z/\sqrt{N}$, where Z = standardised test statistic and N = total number of observations (Field, 2005).

State self-compassion and state self-criticism scores were analysed in two ways. Firstly, Pearson's correlations were run to investigate the relationship between corresponding change scores on the two measures. Zou's method was used to examine whether corresponding correlations for the two conditions differed. Secondly, two separate 2 x 3 mixed model for repeated data analyses were conducted to examine state self-compassion and state self-criticism scores with condition (IVR, MI) as the between-subjects variable and time (pre-intervention, post-intervention, FU) as the within-subjects variable. Unlike other repeated-measure analyses, the mixed models approach does not discard all results from participants who have a single missing measurement (Howell, 2008; Seltman, 2014). As some participants did not complete the FU measures this approach was preferable. A model which assumed sphericity was used. Bayesian information criteria showed this to be a better fit than an unstructured solution which does not make assumptions about the form of covariance matrix. Post-hoc pairwise comparisons were explored for all main effects (regardless of whether the main effect was significant; Motulsky, 2011). The mixed models output required Cohen's d_z effect size to be calculated for within-group analysis using the formula, $t/\sqrt{n}$, where t = t value and n = number of pairs of participants (Lakens, 2013). The effect size was interpreted as small $d_z = 0.2$, medium $d_z = 0.5$, or large $d_z = 0.8$ (Cohen, 1992). Effect size for between-group analysis was not required.

Ease of recall, frequency of practice and mean image vividness scores were investigated using non-parametric tests due to their restricted scales. Difference between the conditions on each measure was analysed using Mann-Whitney tests. To explore effects of the intervention at FU Spearman's correlations were run between these three measures and pure allocentric memory scores, VREQ statements

which indicate the illusion of body ownership and state self-compassion and state self-criticism change scores. Correlations were run for each condition separately and Zou's analysis was used to assess for difference between them.

3: Results

3.1 Baseline Data

Participants in the IVR and MI conditions did not differ on demographic or pre-intervention measures except on the trait self-criticism RS subscale. These scores were higher in the IVR condition than the MI condition. Demographic and trait self-criticism measures did not differ between eligible individuals who did or did not complete the intervention. Data for these series of comparisons are presented in Table 1 and 2 respectively.

3.2 Visuo-Spatial Perspective-Taking

3.2.1 VSPT task data checks

None of the participants had a marked deficit (performed at floor) on the egocentric VSPT task (raw data: $M = 13.80$, $SD = 3.06$, range = 7-16). Results suggested that there was no relationship between the number of correct responses and the time taken to complete a task: egocentric VSPT task, $r(38) = -.08$, $p = .61$; topographical perception task, $r(38) = -.19$, $p = .24$; topographical memory task, $r(38) = .31$, $p = .05$.

3.2.2 Hypothesis 1: Allocentric VSPT ability will be positively associated with change in state self-compassion and may affect state self-criticism. Difference between the conditions is uncertain.

Pure allocentric memory was not related to state self-compassion or state self-criticism change scores in either condition. Zou's analyses showed that there was no difference between the conditions in the strength of their corresponding correlations. Means, standard deviations and correlations with pure allocentric memory for state self-compassion and state self-criticism change scores are presented in Table 3.

(Page left blank intentionally due to formatting)

Table 1

Means, Standard Deviations and Difference Between the IVR and MI Conditions for Demographic and Pre-Intervention Measures

Variable [a]		**IVR (*n* = 20)**			**MI (*n* = 20)**		**Statistic**	***p***	**95% CI**
		Raw data [b]	**Winsorised data**		**Raw data** [b]	**Winsorised data**			
Age		26.90 (8.96)	26.85 (8.83)		24.85 (6.24)	24.75 (5.95)	t = 0.88	.38	[-2.72, 6.92]
Right handed *n* (%)		19 (95)			19 (95)		χ^2 = 0.00	1.00	
Student *n* (%) No		5 (25)			2 (10)		χ^2 = 2.42	.49	
Uni undergrad		6 (30)			7 (35)				
Uni postgrad		9 (45)			10 (50)				
At different Uni					1 (5)				
Trait self-criticism IS		25.50 (3.32)			26.00 (2.92)		t = -0.51	.62	[-2.50, 1.50]
Trait self-criticism HS		5.85 (3.73)	5.80 (3.59)		5.95 (4.94)		t = -0.11	.91	[-2.92, 2.62]
Trait self-criticism RS		20.40 (3.56)			17.45 (4.16)		t = 2.41	.02*	[0.47, 5.43]
State self-compassion		65.70 (26.73)	65.60 (26.55)		59.15 (22.11)	59.00 (21.81)	t = 0.86	.40	[-8.95, 22.15]
State self-criticism		104.20 (20.65)	104.35 (20.09)		100.65 (20.95)	100.65 (20.76)	t = 0.57	.57	[-9.38, 16.78]
Egocentric VSPT correct	R =	13.20 (3.43)		R =	14.40 (2.58)				
	T =	1.20 (1.19)	1.20 (1.19)	T =	0.82 (0.99)		t = 1.11	.27	[-0.32, 1.09]
Egocentric VSPT time	R =	2.61 (1.75)		R =	2.42 (1.38)				
	T =	1.54 (0.50)	1.54 (0.48)	T =	1.51 (0.39)	1.50 (0.36)	t = 0.29	.77	[-0.23, 0.31]
Topographical perception correct	R =	12.50 (2.14)		R =	12.85 (2.32)				
	T =	1.39 (0.77)	1.39 (0.76)	T =	1.22 (0.84)	1.21 (0.82)	t = 0.69	.49	[-0.33, 0.68]
Topographical perception time		14.98 (4.43)	14.99 (4.41)		15.04 (5.25)	15.03 (5.23)	t = -0.02	.98	[-3.13, 3.06]
Topographical memory correct		9.85 (2.66)			9.85 (2.74)		t = 0.00	1.00	[-1.73, 1.73]
Topographical memory time		8.90 (3.22)	8.88 (3.11)		8.89 (3.14)	8.89 (3.08)	t = -0.01	.99	[-1.99, 1.97]
Pure allocentric memory		-2.65 (2.30)			-3.00 (1.95)		t = 0.52	.61	[-1.01, 1.71]

Note. Statistical analysis was performed on winsorised data.

IVR = immersive virtual reality; MI = mental imagery; IS = Inadequate Self subscale, HS = Hated Self subscale, RS = Reassuring Self subscale; VSPT = visuo-spatial perspective-taking; R = raw score; T = transformed score.

[a] age = years; time = seconds.

[b] Mean (*SD*) unless otherwise stated.

*****$p < .05$.

Table 2

Means, Standard Deviations and Difference Between Individuals Who Did or Did Not Complete the Intervention for Demographic and Trait Self-Criticism Measures

Variable		Completed (*n* = 40)	Did not complete (*n* = 21)		Statistic	*p*	95% CI
		Raw data [a b]	Raw data [a]	Winsorised data			
Female *n* (%)		20 (50.0)	15 (71.4)		$\chi^2 = 2.59$	.11	
Right handed *n* (%)		38 (95.0)	18 (85.7)		$\chi^2 = 1.58$	.21	
Student *n* (%)	No	7 (17.5)	9 (42.9)		$\chi^2 = 5.87$	.12	
	Uni undergrad	13 (32.5)	7 (33.3)				
	Uni postgrad	19 (47.5)	5 (23.8)				
	At different Uni	1 (2.5)					
Trait self-criticism IS		25.75 (3.10)	25.95 (3.60)		*t* = -0.23	.82	[-1.97, 1.56]
Trait self-criticism HS		5.90 (4.32)	6.19 (3.34)		*t* = -0.27	.79	[-2.46, 1.87]
Trait self-criticism RS		18.93 (4.10)	18.81 (4.25)	18.86 (4.13)	*t* = 0.06	.95	[-2.15, 2.29]

Note. Statistical analysis was performed on winsorised data.
IS = Inadequate Self subscale, HS = Hated Self subscale, RS = Reassuring Self subscale.
[a] Mean (*SD*) unless otherwise stated.
[b] Winsorising not required.

Table 3

Means, Standard Deviations and Correlations with Pure Allocentric Memory for State Self-Compassion and State Self-Criticism Change Scores

Variable	Change score	IVR			MI			95% CI_z for correlations
			Correlation with PAM			Correlation with PAM		
		M (SD)	*r*	*p*	*M (SD)*	*r*	*p*	
State self-compassion	pre-post	7.30 (17.84)	.03	.89	15.90 (17.47)	.25	.29	[-0.80, 0.42]
	post-FU	0.00 (12.55)	-.32	.21	3.33 (20.55)	-.12	.68	[-0.86, 0.50]
	pre-FU	10.59 (19.44)	-.37	.14	20.33 (23.08)	.03	.91	[-1.01, 0.32]
State self-criticism	pre-post	-15.50 (21.83)	-.35	.13	-24.90 (24.70)	-.09	.71	[-0.83, 0.36]
	post-FU	-4.29 (18.89)	.14	.58	7.40 (22.81)	.19	.49	[-0.73, 0.65]
	pre-FU	-22.35 (23.98)	-.14	.61	-16.00 (22.44)	.12	.68	[-0.91, 0.48]

Note. IVR = immersive virtual reality; MI = mental imagery; PAM = pure allocentric memory; FU = follow-up.

3.3 Embodiment

3.3.1 Question 1: Does sense of embodiment in IVR relate to change in state self-compassion or state self-criticism?

There was no relationship between sense of embodiment and state self-compassion change scores: "As an adult I felt as if the body I saw when I looked in the mirror was my own body", pre-post, $r_s(18) = .26$, $p = .26$; post-FU, $r_s(15) = -.06$, $p = .82$; pre-FU, $r_s(15) = .12$, $p = .64$; "As an adult I had the feeling that I was looking at myself in the mirror rather than looking at someone else", pre-post, $r_s(18) = .32$, $p = .17$; post-FU, $r_s(15) = -.03$, $p = .92$; pre-FU, $r_s(15) = .10$, $p = .70$; "As a child I felt as if the body I saw when I looked in the mirror was my own body", pre-post, $r_s(18) = .14$, $p = .56$; post-FU, $r_s(15) = -.25$, $p = .33$; pre-FU, $r_s(15) = -.15$, $p = .57$.

There was no relationship between sense of embodiment and state self-criticism change scores: "As an adult I felt as if the body I saw when I looked in the mirror was my own body", pre-post, $r_s(18) = -.23$, $p = .34$; post-FU, $r_s(15) = -.31$, $p = .23$; pre-FU, $r_s(15) = -.10$, $p = .70$; "As an adult I had the feeling that I was looking at myself in the mirror rather than looking at someone else", pre-post, $r_s(18) = -.09$, $p = .70$; post-FU, $r_s(15) = -.34$, $p = .18$; pre-FU, $r_s(15) = -.16$, $p = .55$; "As a child I felt as if the body I saw when I looked in the mirror was my own body", pre-post, $r_s(18) = -.05$, $p = .84$; post-FU, $r_s(15) = -.05$, $p = .84$; pre-FU, $r_s(15) = .24$, $p = .36$.

3.3.2 Question 2: Does sense of embodiment in IVR relate to allocentric VSPT ability?

Sense of embodiment was not related to pure allocentric memory: "As an adult I felt as if the body I saw when I looked in the mirror was my own body",

$r_s(18) = .20$, $p = .41$; "As an adult I had the feeling that I was looking at myself in the mirror rather than looking at someone else", $r_s(18) = .30$, $p = .20$; "As a child I felt as if the body I saw when I looked in the mirror was my own body", $r_s(18) = -.03$, $p = .90$.

3.4 State Self-Compassion and State Self-Criticism

Hypothesis 2: The intervention will increase state self-compassion and may affect state self-criticism. Difference in the effect of the conditions is uncertain.

3.4.1 The relationship between state self-compassion and state self-criticism change scores

There was a strong, negative relationship between state self-compassion pre-post change scores and state self-criticism pre-post change scores in both conditions, IVR, $r(18) = -.69$, $p = .001$; MI, $r(18) = -.58$, $p = .008$. These survived Bonferroni correction which set significance to $p < .02$. This indicates that a pre-post intervention increase in state self-compassion was associated with a pre-post intervention decrease in state self-criticism. There was no difference in the strength of these correlations, 95% CI_z [-0.54, 0.30].

Post-FU change scores for state self-compassion and state self-criticism were not related in either condition, IVR, $r(15) = -.32$, $p = .21$; MI, $r(13) = -.24$, $p = .38$. Pre-FU change scores for state self-compassion and state self-criticism also indicated that these variables were not statistically related in either condition, IVR, $r(15) = -.49$, $p = .05$; MI, $r(13) = -.24$, $p = .39$. There was no difference between the conditions in the strength of their post-FU or pre-FU correlations, 95% CI_z [-0.75, 0.59], [-0.79, 0.32], respectively.

3.4.2 Change in state self-compassion and state self-criticism scores

For state self-compassion there was no significant main effect of condition, $F(1, 38.25) = 0.00$, $p = .99$. All post hoc pairwise comparisons were non-significant. Data for these comparisons are presented in Table 4.

There was a significant main effect of time, $F(2, 68.74) = 12.44$, $p < .001$. All post hoc pairwise comparisons for the IVR condition suggested that there was no significant change in scores over time. In the MI condition post hoc pairwise comparisons indicated that state self-compassion scores were lower pre-intervention than post-intervention with a large effect size, and lower pre-intervention than at FU with a large effect size, but no different between post-intervention and FU. All differences survived Bonferroni correction. Data for these comparisons are presented in Table 5.

There was no interaction between time and condition for state self-compassion, $F(2, 68.74) = 1.83$, $p = .17$. Scores for the IVR and MI conditions were similar at each time point. Data for the state self-compassion scores at each time point are presented in Table 6 and Figure 5.

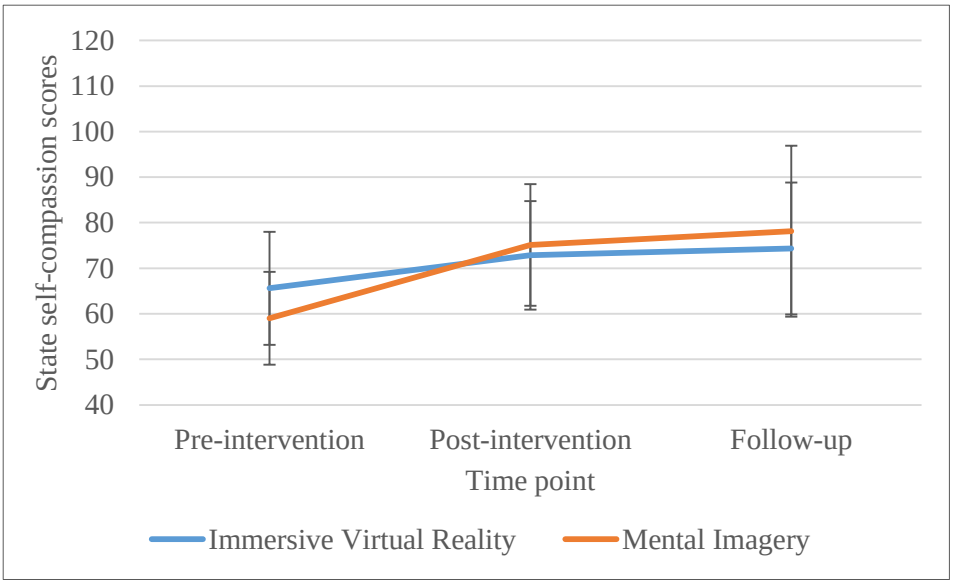


Figure 5. Pre-intervention, post-intervention and two week follow-up state self-compassion scores for the IVR and MI conditions. Error bars represent 95% CIs.

For state self-criticism there was no significant main effect of condition, $F(1, 37.31) = 0.78$, $p = .38$. All post hoc pairwise comparisons were non-significant. Data for these comparisons are presented in Table 4.

There was a significant main effect of time, $F(2, 68.66) = 19.41$, $p < .001$. Post hoc pairwise comparisons indicated that in the IVR condition state self-criticism scores were higher pre-intervention than post-intervention with a medium effect size, and higher pre-intervention than at FU with a large effect size, but no different between post-intervention and FU. Post hoc pairwise comparisons for the MI condition indicated that state self-criticism scores were higher pre-intervention than post-intervention with a large effect size, and higher pre-intervention than at FU with a medium effect size, but no different between post-intervention and FU. All differences survived Bonferroni correction. Data for these comparisons are presented in Table 5.

There was no interaction between time and condition for state self-criticism, $F(2, 68.66) = 1.61$, $p = .21$. Scores for the IVR and MI conditions were similar at

each time point. Data for the state self-criticism scores at each time point are presented in Table 6 and Figure 6.

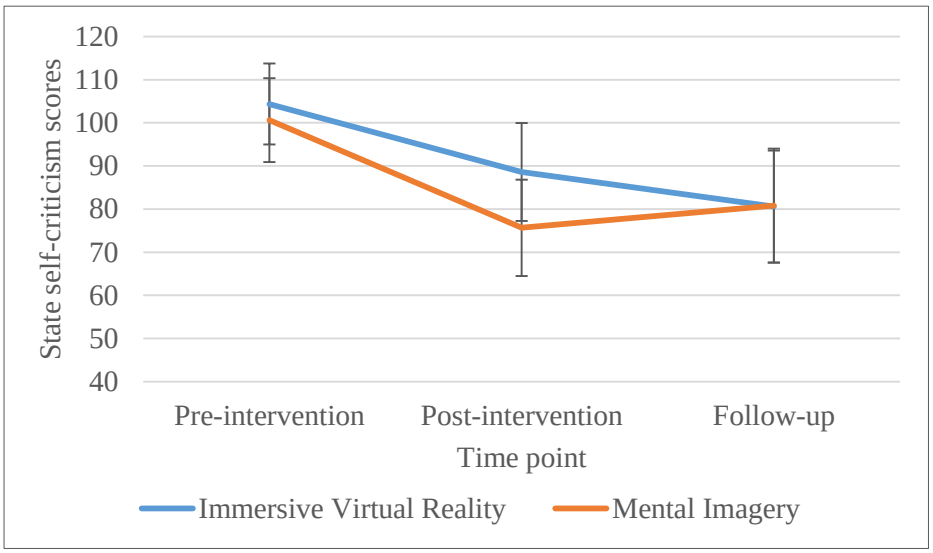


Figure 6. Pre-intervention, post-intervention and two week follow-up state self-criticism scores for the IVR and MI conditions. Error bars represent 95% CIs.

Table 4

Post Hoc Pairwise Comparisons for the Main Effect of Condition for State Self-Compassion and State Self-Criticism

Variable	**Time point**	***M* difference *(SE)* between IVR and MI conditions**	***p***	**95% CI**
State self-compassion	pre	6.60 (8.59)	.45	[-10.64, 23.84]
	post	-2.25 (8.59)	.79	[-19.49, 14.99]
	FUMA	-4.54 (8.95)	.61	[-22.45, 13.37]
State self-criticism	pre	3.70 (7.33)	.62	[-10.90, 18.30]
	post	12.95 (7.33)	.08	[-1.65, 27.55]
	FUMA	-0.64 (7.92)	.94	[-16.40, 15.12]

Note. IVR = immersive virtual reality; MI = mental imagery; FUMA = follow-up mixed model adjusted.

Table 5

Post Hoc Pairwise Comparisons for the Main Effect of Time for State Self-Compassion and State Self-Criticism

Condition	Variable	Time period	*M* difference *(SE)* over time	*p* [a]	d_z [b]	95% CI
IVR	State self-compassion	pre to post	-7.25 (4.16)	.09		[-15.54, 1.04]
		post to FUMA	-1.57 (4.41)	.72		[-10.37, 7.23]
		pre to FUMA	-8.82 (4.41)	.05		[-17.62, -0.02]
MI		pre to post	-16.10 (4.16)	< .001**	-0.87	[-24.39, -7.81]
		post to FUMA	-3.86 (4.63)	.41		[-13.08, 5.37]
		pre to FUMA	-19.96 (4.63)	< .001**	-0.96	[-29.18, -10.73]
IVR	State self-criticism	pre to post	15.75 (5.19)	.003**	0.68	[5.40, 26.10]
		post to FUMA	6.52 (5.48)	.24		[-4.42, 17.46]
		pre to FUMA	22.27 (5.48)	< .001**	0.91	[11.34, 33.21]
MI		pre to post	25.00 (5.19)	< .001**	1.08	[14.65, 35.35]
		post to FUMA	-7.07 (5.73)	.22		[-18.50, 4.36]
		pre to FUMA	17.93 (5.73)	.003**	0.70	[6.50, 29.36]

Note. IVR = immersive virtual reality; MI = mental imagery; FUMA = follow-up mixed model adjusted.
[a] Significance level: $p < .05$ or Bonferroni correction $p < .02$.
[b] Cohen's d_z effect size: small = 0.2, medium = 0.5, large = 0.8.
**Significant at $p < .02$.

Table 6

State Self-Compassion and State Self-Criticism at Each Time Point

Variable	Time point	IVR			MI		
		M (SD)	*SE*	95% CI	*M (SD)*	*SE*	95% CI
State self-compassion	pre	65.60 (26.55)	6.08	[53.41, 77.79]	59.00 (21.81)	6.08	[46.81, 71.19]
	post	72.85 (25.46)	6.08	[60.66, 85.04]	75.10 (28.50)	6.08	[62.91, 87.29]
	FU	74.35 (28.20)			78.13 (33.93)		
	FUMA	74.42	6.25	[61.90, 86.94]	78.96	6.41	[66.15, 91.77]
State self-criticism	pre	104.35 (20.09)	5.18	[94.03, 114.67]	100.65 (20.76)	5.18	[90.33, 110.97]
	post	88.60 (24.29)	5.18	[78.28, 98.92]	75.65 (23.92)	5.18	[65.33, 85.97]
	FU	80.59 (25.30)			80.80 (23.85)		
	FUMA	82.08	5.48	[71.18, 92.98]	82.72	5.73	[71.34, 94.10]

Note. IVR = immersive virtual reality; MI = mental imagery; FU = follow-up; FUMA = follow-up mixed model adjusted.

3.5 Experience of the Intervention

Question 3: Are the interventions experienced differently?

Table F1 in Appendix F presents the medians, interquartile ranges and difference between corresponding statements for the VREQ (completed in the IVR condition) and MIEQ (completed in the MI condition). Participants in the IVR condition were more concerned about forgetting their lines and found it easier to recognise themselves in the adult avatar's voice. They felt less comforted and less reassured by the adult avatar and more critical of them. Differences between the conditions were of medium effect size, except for feeling comforted where effect size was large. Results for feeling comforted and reassured survived Bonferroni correction. There were no other differences between the conditions.

3.6 Effects at Two Week Follow-Up

3.6.1 Difference between the conditions

There was no difference between the conditions on measures of ease of recall, frequency of practice or image vividness. Table 7 presents the medians, interquartile ranges and difference between the conditions for the above measures.

Table 7

Medians, Interquartile Ranges and Difference Between the Conditions for Ease of Recall, Frequency of Practice and Image Vividness

Variable	**IVR *Mdn (IQR)***	**MI *Mdn (IQR)***	***U***	***p***
Ease of recall	4.00 (4.00-4.50)	4.00 (3.00-5.00)	131.50	.86
Frequency of practice	4.00 (4.00-4.50)	4.00 (3.00-5.00)	124.00	.65
Image vividness	2.50 (1.83-2.83)	2.33 (1.62-3.00)	125.50	.70

Note. IVR = immersive virtual reality; MI = mental imagery.

3.6.2 Question 4: Does frequency of practice, experiencing the image vividly or recalling it with ease for two weeks after the intervention relate to allocentric VSPT ability or sense of embodiment in IVR? Does this differ between the conditions?

There was a strong, positive relationship between ease of recall and pure allocentric memory in the MI condition suggesting that recall is easier for those with greater allocentric VSPT ability. This survived Bonferroni correction. This relationship was not present in the IVR condition and Zou's analysis showed a significant difference in the strength of this relationship between the conditions. Frequency of practice and image vividness were not related to pure allocentric memory in either condition. None of these FU measures were related to sense of embodiment. All other Zou's analyses were non-significant. Data for these correlations are presented in Table 8.

3.6.3 Hypothesis 3: Frequency of practice, experiencing the image vividly or recalling it with ease for two weeks after the intervention will lead to greater increase in state self-compassion and may affect state self-criticism. Difference in the effect of the conditions is uncertain.

Correlations for ease of recall, frequency of practice and image vividness with pre-FU and post-FU state self-compassion and self-criticism change scores in each condition found one significant association. This was a strong, positive relationship between image vividness and the state self-criticism pre-FU change score in the IVR condition. This suggests that reduction in state self-criticism between pre-intervention and FU was associated with experiencing images of the intervention vividly. The finding did not survive Bonferroni correction. All Zou's analyses were non-significant. Data for these correlations are presented in Table 8.

Table 8

Correlations for Ease of Recall, Frequency of Practice and Image Vividness with Pure Allocentric Memory, Sense of Embodiment and State Self-Compassion and State Self-Criticism Change Scores

Variables			**IVR**		**MI**		**95% CI$_z$**
FU measure	**Other measure**		**r_s (15)**	**p [a]**	**r_s (13)**	**p [a]**	
Ease of recall	Pure allocentric memory		.01	.97	.67 [b]	.004**	[-1.18, -0.03] †
	VREQ	As an adult I felt as if the body I saw when I looked in the mirror was my own body	-.12	.64			
		As an adult I had the feeling that I was looking at myself in the mirror rather than looking at someone else	.23	.37			
		As a child I felt as if the body I saw when I looked in the mirror was my own body	-.19	.46			
	Self-compassion	post-FU change score	.15	.58	-.39	.15	[-0.32, 1.01]
		pre-FU change score	.10	.70	-.51	.05	[-0.11, 1.15]
	Self-criticism	post-FU change score	-.37	.15	.41	.13	[-0.92, 0.40]
		pre-FU change score	-.25	.33	.21	.46	[-1.06, 0.29]
Frequency of practice	Pure allocentric memory		.18	.48	.33	.21	[-0.79, 0.55]
	VREQ	As an adult I felt as if the body I saw when I looked in the mirror was my own body	-.16	.53			
		As an adult I had the feeling that I was looking at myself in the mirror rather than looking at someone else	-.15	.57			
		As a child I felt as if the body I saw when I looked in the mirror was my own body	-.38	.13			
	Self-compassion	post-FU change score	.04	.88	.26	.35	[-0.58, 0.74]
		pre-FU change score	.19	.46	.43	.11	[-0.85, 0.44]
	Self-criticism	post-FU change score	-.19	.46	.11	.70	[-0.76, 0.64]
		pre-FU change score	-.33	.20	-.30	.28	[-0.69, 0.62]

Table continues

Variables			IVR		MI		95% CI_z
FU measure	**Other measure**		r_s **(15)**	p [a]	r_s **(13)**	p [a]	
Image vividness	Pure allocentric memory		-.28	.28	-.13	.63	[-0.82, 0.55]
	VREQ	As an adult I felt as if the body I saw when I looked in the mirror was my own body	.18	.50			
		As an adult I had the feeling that I was looking at myself in the mirror rather than looking at someone else	.00	1.00			
		As a child I felt as if the body I saw when I looked in the mirror was my own body	.36	.16			
	Self-compassion	post-FU change score	-.12	.64	.21	.45	[-0.97, 0.41]
		pre-FU change score	-.13	.61	-.00	.99	[-0.81, 0.59]
	Self-criticism	post-FU change score	.33	.20	.14	.61	[-0.58, 0.78]
		pre-FU change score	.55	.02*****	.32	.25	[-0.37, 0.84]

Note. IVR = immersive virtual reality; MI = mental imagery; FU = follow-up; VREQ = Virtual Reality Experience Questionnaire.

[a] Significance level: $p < .05$ or Bonferroni correction $p < .02$.

[b] $r_s(14)$.

***** Significant at $p < .05$.

****** Significant at $p < .02$.

† Significant difference between the IVR and MI conditions in the strength of the association.

4: Discussion

4.1 Main Findings

This study explored factors which may affect the ability to cultivate self-compassion in self-critical individuals using an IVR or MI intervention and examined change in state self-compassion and self-criticism following the intervention. Contrary to the hypothesis, correlational analysis indicated that scores on the allocentric VSPT measure were statistically unrelated to change in state self-compassion and self-criticism scores. Likewise, correlational analysis suggested that scores on the sense of embodiment measure were statistically unrelated to change in state self-compassion and self-criticism scores in the IVR condition. Scores on the state self-compassion measure only increased in the MI condition but scores on the state self-criticism measure reduced in both conditions. Interestingly, the results indicated that the IVR intervention was experienced more negatively than the MI intervention. Correlational analysis also suggested that at two week FU greater ease of recall was related to greater allocentric VSPT ability in the MI condition and greater image vividness was related to pre-FU reduction in state self-criticism scores in the IVR condition.

4.2 Statistical Considerations

Extensive analysis increases the risk of Type I error and could account for difference between the conditions on the FSCRS Reassuring Self subscale. However, this approach to the analysis seemed appropriate given the exploratory nature of the study. For the same reason both corrected and uncorrected results for multiple comparisons were reported. This avoided missing potentially interesting findings whilst highlighting possible false positives (Linacre, n.d.). All findings are

discussed but those that did not survive correction should be regarded with greater caution. Caution in interpreting the findings is also warranted given that the study used a small sample and the majority of the analysis was correlational which does not imply a causal relationship between the variables.

It was hoped that common effect size indices could be used to aid comparison and generalisability. However, mixed model analysis limited within-group effect size to Cohen's d_z. As this takes correlation between measures into account Dunlap, Cortina, Vaslow, and Burke (1996) argue that it overestimates *true* effect size, preventing generalisation. Furthermore, it does not allow Hedge's *g* adjustment for small sample size (Lakens, 2013).

4.3 Interpretation of Findings

4.3.1 Allocentric VSPT ability

The suggested lack of association between allocentric VSPT ability and change in state self-compassion and self-criticism seems counter to other findings and the theoretical underpinnings of self-compassionate MI (Gilbert & Irons, 2005; Hamilton et al., 2014). This may relate to the requirements of the Four Mountains task. This was chosen for its feasibility within the constraints on the study's resources. However, it assessed allocentric processing of global landscape features (Hartley et al., 2007) whereas social perspective-taking in the intervention may have required allocentric processing of localised features, for example features of the child's disposition (Srinivasan & Gupta, 2011).

An interesting finding was that individuals in the MI condition who found it harder to undertake global allocentric VSPT found it more difficult to recall the scenario across two weeks post-intervention. The suggested relationship between

these variables may be explained by the requirement to generate a global representation of another person with only audio support in the MI condition. Those who found this hard may not have created a sufficiently strong memory representation of the compassionate other for it to be easily retrieved post-intervention. In contrast IVR immersed individuals in a rich, sensory experience with concrete stimuli. This may have generated a salient, distinctive memory representation of an alternative perspective which increased its accessibility (Brewin, 2006). Recall is critical to the long-term efficacy of self-compassionate interventions (Brewin, 1989; Gilbert et al., 2006). Therefore, this indicated relationship could suggest that IVR may be more efficacious than MI at cultivating self-compassion in individuals who struggle with global allocentric VSPT. However, whilst IVR could have removed the effect of VSPT ability on recall, recall was unrelated to change in state self-compassion and self-criticism and IVR participants' mean level of state self-compassion did not increase. This requires further consideration.

4.3.2 Sense of embodiment

Correlational analysis indicated that sense of embodiment in IVR was also unrelated to change in state self-compassion and self-criticism and was unrelated to frequency of practice, image vividness and ease of recall. The null result for change in state self-compassion and self-criticism replicates Falconer et al.'s (2014) finding. The psychometric properties of the VREQ are not established therefore it is possible that it does not have adequate sensitivity and specificity to detect differences.

Results also indicated that there was no association betweem sense of embodiment and allocentric VSPT ability. It is possible that these variables may not be statistically related as the participant was immersed in the avatar's body from a

first person perspective. Perspective-taking in this circumstance may have required embodied allocentric VSPT skills as the avatar's body provided a frame of reference for visuo-spatial computation (Becchio, Del Giudice, Dal Monte, Latini-Corazzini, & Pia, 2013; Tversky & Hard, 2009). However, the VSPT task demanded disembodied processing, taking an alternative viewpoint in the absence of another person in the scene. This is distinct to embodied processing (Vastano, Sulpizio, Steinisch, Comani, & Committeri, 2014) and therefore perhaps unrelated to the VSPT ability required by the intervention.

4.3.3 Change in state self compassion and state self-criticism

The results indicated that mean scores for state self-compassion and state self-criticism in the IVR and MI conditions were not significantly different and that difference in the scores over time did not depend on the condition. However, mean state self-compassion scores in the MI condition were higher post-intervention and at FU compared to pre-intervention. VREQ and MIEQ data suggest that participants in the IVR condition felt less comforted and reassured by the adult and more critical of them than participants in the MI condition. The IVR avatars were standardised, had no eye movement, neutral facial expressions and on occasion unnatural body movements. As such, they may not have been experienced as compassionate in their appearance and lacked the personal dimension of an imagined other (Gilbert, 2009a; Gilbert & Procter, 2006). Participants in the IVR condition were also more concerned about forgetting their lines and more easily recognised themselves in the adult's voice. These participants said their lines aloud overheard by the researchers then heard them back, potentially introducing a confounding level of social anxiety. In the MI intervention these elements were not present and participants had more

control over the content of their images. Consequently, although change in state self-compassion scores over time were not statistically determined by the condition it seems possible that aspects of the MI intervention could have contributed to a somewhat more compassion-nurturing experience (Spagnolli et al., 2014; Zhang et al., 2014). Based on the above considerations, it may be that the efficacy of the IVR intervention could be increased by addressing aspects of the intervention that were negatively experienced.

State self-criticism reduced post-intervention in both conditions equally. Gilbert's social mentality theory suggests that the overall experience of giving and receiving compassion may have been powerful enough to reduce activation of the threat system despite differences in the experience of the interventions (Gilbert, 2005, 2014b). Furthermore, in the IVR condition greater image vividness was associated with greater pre-FU reduction in state self-criticism. This fits with evidence that therapeutic benefit is derived from the ability to visualise vivid self-compassionate images (Kelly et al., 2010). The enhancing effect of concrete stimuli on image vividness may account for this effect only occurring in the IVR condition (Campos, Gómez-Juncal, & Pérez-Fabello, 2008).

Despite differences in mean affect change, in both conditions pre-post intervention increase in state self-compassion was strongly associated with pre-post intervention reduction in state self-criticism. This indicates that there is a dynamic relationship between the two constructs. This is counter to previous suggestion that state self-compassion and state self-criticism are experienced orthogonally (Falconer et al., 2014, 2015). It therefore seems that further investigation into the relationship between these constructs is required. However, the current findings add to evidence

that state self-compassion and self-criticism are sensitive to situational factors (Breines & Chen, 2013; Falconer et al., 2015).

4.3.4 Additional effects at FU

Lack of change in state self-compassion and self-criticism between post-intervention and FU may be attributable to the limited ability of a one-off intervention to induce continued change (Kaplan et al., 2015). The benefit of frequently practicing imagining an intervention found by other studies (Kaplan et al., 2015; Neff & Pommier, 2013) was not replicated here. Despite the findings for ease of recall and image vividness discussed above, the median score for these measures and frequency of practice did not differ between the conditions. However, the interquartile range for these measures was greater in the MI condition than the IVR condition, particularly for ease of recall and frequency of practice, suggesting greater variability in the data for the former condition. Therefore, the null finding for the relationship between these FU measures and state self-compassion and self-criticism change scores could relate to the sensitivity of the FU measures and the limited ability of non-parametric statistics to express subtle differences between the datasets (Field, 2005).

4.4 Limitations

A major limitation of the current study was the assessment measures. Firstly, validated measures for some of the constructs investigated in the current study have yet to be developed. Consequently, measures with unknown psychometric properties had to be used to assess the sense of embodiment and experience of the intervention, frequency of practice and ease of recall. Therefore, the reliability and validity of

results concerning these measures is questionable. This could hinder the ability of future research to build on the current study. Secondly, the study relied on subjective self-report measures which can suffer from social desirability response bias. In order to obtain informed consent it was not possible to ask participants to complete the measures without prior knowledge of the study. So to attempt to reduce this bias the measures were administered online in the absence of a researcher and participants were reminded that data were stored anonymously. Supplementing the measures with psychophysiological assessments was perhaps beyond the scope of the current study. However, inclusion of a social desirability assessment such as the Marlowe-Crowne Social Desirability scale (Strahan & Gerbasi, 1972) would have been possible. Results of this scale could have been correlated with the other measures to assess biased responding. Participants' effort and attention may have differed if the measures were administered manually and may also have differed between measures completed in the lab and those completed at a location of their choosing. Furthermore, Opinio, Version 6.8 (2014) prevented counter-balancing which may have led to order effects. However, for this study, the efficiency, standardisation and removal of researcher bias offered by online administration was considered preferable to manualised administration. Completing the screening and FU survey at a location and time the participant chose seemed the most feasible method of collecting this data within the timeframe and resources available for the study.

It was noted that participants appeared to have different motivations for taking part in the study. Some mentioned that they were keen to experience IVR whilst others were interested in compassion. Immersion in IVR was novel for most participants and variation was observed in their reaction to the experience. Some appeared excited and focused on experiencing the technology rather than the

compassionate elements of the scenario. Others appeared somewhat nervous. These factors may have affected engagement with the intervention. Additionally, the extent to which participants engaged with the MI scenario is unknown. This is a well-recognised confound with research showing that lower levels of engagement are associated with reduced efficacy of MI therapeutic interventions (Odou & Vella-Brodrick, 2013).

The study was mainly advertised at a university and the majority of participants were well educated university students. Ethnicity data were not collected however, it seemed that in respect of this and other socio-demographic variables the sample was not overly representative of the city in which it was located. The study investigated a healthy population who reported no prior experience of mental health difficulties. These factors limit the generalisability of the findings.

The majority of analysis was correlational which does not provide information about causality. Therefore, the direction of influence (i.e., which variable is affecting which) and the role of other variables in generating associations are unknown. Extensive analysis may have increased the risk of Type I error. Additionally, the study had a small sample size and was slightly underpowered which is likely to have increased the chance of Type II error. It is hoped that future research will build on this exploratory study using focussed and refined experimental approaches which will reduce the above limitations.

4.5 Future Research

Future research should initially focus on replicating the current study to confirm the results. It would be beneficial for this to be carried out using a larger sample size to increase the power of the analyses. Replication with a more diverse

socio-demographic sample would also be worthwhile to extend the generalisability of the results (see Stellar, Manzo, Kraus, & Keltner, 2012, for evidence of class-based differences in trait self-compassion). Following this, examining the efficacy of these interventions with clinical populations in clinical settings is warranted to establish their potential as treatments in ecologically valid contexts.

Conducting more detailed quantitative and qualitative assessments of participants' experience of the interventions may provide helpful information about ways to improve the efficacy of these therapeutic tools. Prior to carrying out the intervention this could include assessment of participants' expectations, prior knowledge and experience of using IVR and MI and level of anxiety. Post-intervention assessment could include participants' experience of the environment, the avatars/people imagined, managing the equipment and the flow of the intervention, along with suggestions for improvement. Data from the current study indicates that the investigated IVR intervention may benefit from improving the comforting and reassuring nature of the adult avatar and reducing concern about remembering lines. A step towards addressing the former concern could be programming the avatars to have changeable facial expressions and blinking eyes. The latter concern may be reduced by hearing each line directly before it is to be spoken. However, careful assessment of the effect of such changes would be needed.

In order for the MI intervention to be an analogue of the IVR intervention it included embodiment exercises and a detailed description of the child. This may have reduced its evocativeness in comparison to typical compassionate-self MI exercises. These only include components directly relevant to generating mental images and are less prescriptive as it is considered therapeutically beneficial for participants to generate personally meaningful and relevant images (Gilbert &

Procter, 2006; Zhang et al., 2014). Comparing the existing scenarios with altered versions and typical compassionate-self MI exercises which are currently used in clinical practice may help to further understanding of their effective components.

Findings from the current study could be confirmed and extended by using additional and validated measures. These could include an embodied allocentric VSPT task, behavioural and physiological measures of compassion (Rockliff et al., 2008) and event related potentials to measure embodiment (González-Franco, Peck, Rodríguez-Fornells, & Slater, 2014). It is possible that a daily report of practice, vividness and ease of recall could increase the validity of these FU measures. None of the participants reported simulator sickness however future studies may benefit from quantitative assessment of this and other potential confounds such as an individual's ability to generate detailed mental images (Pearson, Rademaker, & Tong, 2011).

The current study could also be extended by the addition of a control condition to account for non-specific effects. Longer term follow-up of analogue IVR and MI interventions designed to cultivate self-compassion would enable greater exploration of the comparative long-term effectiveness of these treatment tools. Findings indicate that the assessment of the long-term relationship between recall and allocentric VSPT ability may be particularly worthwhile exploring.

4.6 Clinical Implications

The current study suggests that in relation to the scenarios investigated, aspects of the MI condition may have contributed to an increase in state self-compassion in highly self-critical individuals. However, findings also indicate that

many factors may need to be considered when designing and choosing an intervention and selecting an appropriate treatment tool.

If use of the MI scenario investigated here or a similar compassionate-self MI scenario which involves allocentric perspective-taking is being considered, therapists may benefit from assessing the client's global allocentric VSPT ability. If this ability is impaired in the client it may hinder their ability to generate a sufficiently strong memory representation to recall the scenario. Ensuring that imagery is distinctive and incorporates rich sensory detail may offer some help with this problem (Brewin, 2006). Alternatively, the individual may benefit from carrying out the intervention using IVR. Additionally, if individuals are struggling to recall MI interventions which involve allocentric perspective-taking a global allocentric VSPT task could be administered to help determine the causes of the difficulty.

If considering using the IVR intervention investigated in the current study it may be important to note that some individuals found aspects of this intervention to be aversive. In line with Gilbert and Irons (2005), this finding suggests that providing a compassion-nurturing environment may be an essential component of an intervention designed to cultivate self-compassion. Therefore, the client's response to potentially aversive factors such as the avatars appearance, remembering lines and hearing their own voice should be taken into account. (Effective ways to address such factors in IVR can only be established through further research). Findings tentatively suggest that the IVR intervention investigated may be indicated as a tool to generate vivid images which may help to reduce state self-criticism. The increasing affordability and commercial availability of IVR technology supports its potential for use in clinical settings. As such, therapists will need to consider which

treatment tool will be most efficacious for the client, a decision which can become better informed with further development of the evidence-base.

4.7 Conclusion

Rather than allocentric VSPT or sense of embodiment it seems that the experience of the IVR and MI interventions may have contributed to change in state self-compassion and self-criticism. Overall, the IVR and MI interventions were comparable in their ability to reduce state self-criticism however, state self-compassion only increased following the MI intervention. It may be possible that addressing aspects of the IVR intervention that were negatively experienced could increase its efficacy. Participants in the MI condition who had lower global allocentric VSPT ability found it harder to recall the intervention. Participants in the IVR condition who experienced vivid images of the intervention reported a greater pre-FU reduction in state self-criticism. This perhaps suggests that, with development, IVR may be indicated as a tool for cultivating state self-compassion in individuals who may struggle with global allocentric VSPT or to support the generation of vivid images to help reduce state self-criticism.

References

Ahn, S. J., Le, A. M. T., & Bailenson, J. (2013). The effect of embodied experiences on self-other merging, attitude, and helping behavior. *Media Psychology, 16*(1), 7–38. doi:10.1080/15213269.2012.755877

Banakou, D., Groten, R., & Slater, M. (2013). Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(31), 12846–12851. doi:10.1073/pnas.1306779110

Barnard, L. K., & Curry, J. F. (2011). Self-compassion: Conceptualizations, correlates, and interventions. *Review of General Psychology, 15*(4), 289–303. doi:10.1037/a0025754

Becchio, C., Del Giudice, M., Dal Monte, O., Latini-Corazzini, L., & Pia, L. (2013). In your place: Neuropsychological evidence for altercentric remapping in embodied perspective taking. *Social Cognitive and Affective Neuroscience, 8,* 165–170. doi:10.1093/scan/nsr083

Bird, C. M., Capponi, C., King, J. A., Doeller, C. F., & Burgess, N. (2010). Establishing the boundaries: The hippocampal contribution to imagining scenes. *The Journal of Neuroscience, 30*(35), 11688–11695. doi:10.1523/JNEUROSCI.0723-10.2010

Blanke, O. (2012). Multisensory brain mechanisms of bodily self-consciousness. *Nature Reviews Neuroscience, 13,* 556–571. doi:10.1038/nrn3292

Breines, J. G., & Chen, S. (2013). Activating the inner caregiver: The role of support-giving schemas in increasing state self-compassion. *Journal of Experimental Social Psychology*, *49*(1), 58–64. doi:10.1016/j.jesp.2012.07.015

Brewin, C. R. (1989). Cognitive change processes in psychotherapy. *Psychological Review*, *96*(3), 379–394. Retrieved from http://discovery.ucl.ac.uk/125403/

Brewin, C. R. (2006). Understanding cognitive behaviour therapy: A retrieval competition account. *Behaviour Research and Therapy*, *44*, 765–784. doi:10.1016/j.brat.2006.02.005

Campos, A., Gómez-Juncal, R., & Pérez-Fabello, M. J. (2008). Experience in imagery and imagery vividness. *Imagination, Cognition and Personality*, *27*(4), 337–348. doi:10.2190/IC.27.4.d

Clements-Stephens, A. M., Vasiljevic, K., Murray, A. J., & Shelton, A. L. (2013). The role of potential agents in making spatial perspective taking social. *Frontiers in Human Neuroscience*, *7*, 1–11. doi:10.3389/fnhum.2013.00497

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *122*(1), 155–159.

Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, *1*(2), 170–177. doi:10.1037/1082-989X.1.2.170

Eisenberg, N., Wentzel, N. M., & Harris, J. D. (1998). The role of emotionality and regulation in empathy-related responding. *School Psychology Review*, *27*(4), 506–521. Retrieved from http://ezproxy.umsl.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eft&AN=507616737&site=ehost-live&scope=site

Falconer, C. J., King, J. A., & Brewin, C. R. (2015). Demonstrating mood repair with a situation-based measure of self-compassion and self-criticism. *Psychology and Psychotherapy: Theory, Research and Practice*, 1–15. doi:10.1111/papt.12056

Falconer, C. J., Slater, M., Rovira, A., King, J. A., Gilbert, P., Antley, A., & Brewin, C. R. (2014). Embodying compassion: A virtual reality paradigm for

overcoming excessive self-criticism. *PLoS ONE*, *9*(11). doi:e111933. doi:10.1371/journal.pone.0111933

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage Publications Ltd. doi:10.1016/j.landurbplan.2008.06.008

Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing. *Neuron*, *50*, 531–534. doi:10.1016/j.neuron.2006.05.001

Frith, U., & De Vignemont, F. (2005). Egocentrism, allocentrism, and asperger syndrome. *Consciousness and Cognition*, *14*, 719–738. doi:10.1016/j.concog.2005.04.006

Gaesser, B., & Schacter, D. L. (2014). Episodic simulation and episodic memory can increase intentions to help others. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(12), 4415–4420. doi:10.1073/pnas.1402461111

Gamito, P., Oliveira, J., Rosa, P., Morais, D., Duarte, N., Oliveira, S., & Saraiva, T. (2010). PTSD elderly war veterans: A clinical controlled pilot study. *Cyberpsychology, Behavior and Social Networking*, *13*(1), 43–48. doi:10.1089=cyber.2009.0237

Gardner, M. R., Sorhus, I., Edmonds, C. J., & Potts, R. (2012). Sex differences in components of imagined perspective transformation. *Acta Psychologica*, *140*(1), 1–6. doi:10.1016/j.actpsy.2012.02.002

Germer, C. K., & Neff, K. D. (2013). Self-compassion in clinical practice. *Journal of Clinical Psychology*, *69*(8), 856–867. doi:10.1002/jclp.22021

Ghosh, D., & Vogt, A. (2012). Outliers: An evaluation of methodologies. In *Joint Statistical Meetings* (pp. 3455–3460).

Gilbert, P. (2000). Social mentalities: Internal "social" conflicts and the role of inner warmth and compassion in cognitive therapy. In P. Gilbert & K. G. Bailey (Eds.), *Genes on the couch: Explorations in evolutionary psychotherapy* (pp. 118–150). Hove: Brunner-Routledge.

Gilbert, P. (2005). Compassion and cruelty: A biopsychosocial approach. In P. Gilbert (Ed.), *Compassion: Conceptualisations, research and use in psychotherapy* (pp. 9–74). East Sussex: Routledge.

Gilbert, P. (2009a). Introducing compassion-focused therapy. *Advances in Psychiatric Treatment*, *15*, 199–208. doi:10.1192/apt.bp.107.005264

Gilbert, P. (2009b). The nature and basis for compassion focused therapy. *Hellenic Journal of Psychology*, *6*, 273–291.

Gilbert, P. (2014a). Compassion-focused therapy: Preface and introduction for special section. *British Journal of Clinical Psychology*, *53*, 1–5. doi:10.1111/bjc.12045

Gilbert, P. (2014b). The origins and nature of compassion focused therapy. *British Journal of Clinical Psychology*, *53*, 6–41. doi:10.1111/bjc.12043

Gilbert, P., Baldwin, M. W., Irons, C., Baccus, J. R., & Palmer, M. (2006). Self-criticism and self-warmth: An imagery study exploring their relation to depression. *Journal of Cognitive Psychotherapy*, *20*(2), 183–200. doi:10.1891/jcop.20.2.183

Gilbert, P., & Choden. (2013). *Mindful compassion*. London: Constable & Robinson Ltd.

Gilbert, P., Clarke, M., Hempel, S., Miles, J. N. V, & Irons, C. (2004). Criticizing and reassuring oneself: An exploration of forms, styles and reasons in female students. *The British Journal of Clinical Psychology*, *43*, 31–50. doi:10.1348/014466504772812959

Gilbert, P., & Irons, C. (2004). A pilot exploration of the use of compassionate images in a group of self-critical people. *Memory*, *12*(4), 507–516. doi:10.1080/09658210444000115

Gilbert, P., & Irons, C. (2005). Focused therapies and compassionate mind training for shame and self-attacking. In P. Gilbert (Ed.), *Compassion: Conceptualisations, research and use in psychotherapy* (pp. 263–325). East Sussex: Routledge.

Gilbert, P., & Procter, S. (2006). Compassionate mind training for people with high shame and self-criticism: Overview and pilot study of a group therapy approach. *Clinical Psychology and Psychotherapy*, *13*, 353–379. doi:10.1002/cpp

Gillath, O., McCall, C., Shaver, P. R., & Blascovich, J. (2008). What can virtual reality teach us about prosocial tendencies in real and virtual environments? *Media Psychology*, *11*(2), 259–282. doi:10.1080/15213260801906489

González-Franco, M., Peck, T. C., Rodríguez-Fornells, A., & Slater, M. (2014). A threat to a virtual hand elicits motor cortex activation. *Experimental Brain Research*, *232*, 875–887. doi:10.1007/s00221-013-3800-1

González-Franco, M., Pérez-Marcos, D., Spanlang, B., & Slater, M. (2010). The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment. In *Proceedings - IEEE Virtual Reality* (pp. 111–114). doi:10.1109/VR.2010.5444805

Gregg, L., & Tarrier, N. (2007). Virtual reality in mental health: A review of the literature. *Social Psychiatry and Psychiatric Epidemiology*, *42*, 343–354. doi:10.1007/s00127-007-0173-4

Hackmann, A., & Holmes, E. A. (2004). Reflecting on imagery: A clinical perspective and overview of the special issue of memory on mental imagery and memory in psychopathology. *Memory*, *12*(4), 389–402. doi:10.1080/09658210444000133

Hamilton, A. F., Brindley, R., & Frith, U. (2009). Visual perspective taking impairment in children with autistic spectrum disorder. *Cognition*, *113*(1), 37–44. doi:10.1016/j.cognition.2009.07.007

Hamilton, A. F., Kessler, K., & Creem-Regehr, S. H. (2014). Perspective taking: Building a neurocognitive framework for integrating the "social" and the "spatial." *Frontiers in Human Neuroscience*, *8*. doi:10.3389/fnhum.2014.00403

Hartley, T., Bird, C. M., Chan, D., Cipolotti, L., Husain, M., Varga-Khadem, F., & Burgess, N. (2007). The hippocampus is required for short-term topographical memory in humans. *Hippocampus*, *17*, 34–38. doi:10.1002/hipo.20240

Holden, A. (n.d.). A comparison of virtual reality and mental imagery scenarios to promote self-compassion and reduce shame and self-criticism. *Unpublished*.

Howell, D. C. (2008). *Mixed models for repeated (longitudinal) data*. Retrieved January 27, 2015, from http://www.uvm.edu/~dhowell/StatPages/More_Stuff/Missing_Data/Mixed Models for Repeated Measures.pdf

Kaplan, U., & Epstein, G. N. (2012). Psychophysiological coherence as a function of mental imagery practice. *Imagination, Cognition and Personality*, *31*(4), 297–312. doi:10.2190/IC.31.4.d

Kaplan, U., Epstein, G. N., & Sullivan-Smith, A. (2015). Microdevelopment of daily well-being through mental imagery practice. *Imagination, Cognition and Personality*, *34*(1), 73–96.

Kelly, A. C., Zuroff, D. C., Foa, C. L., & Gilbert, P. (2010). Who benefits from training in self-compassionate self-regulation? A study of smoking reduction. *Journal of Social and Clinical Psychology*, *29*(7), 727–755. doi:10.1521/jscp.2010.29.7.727

Kelly, A. C., Zuroff, D. C., & Shapira, L. B. (2009). Soothing oneself and resisting self-attacks: The treatment of two intrapersonal deficits in depression vulnerability. *Cognitive Therapy and Research*, *33*, 301–313. doi:10.1007/s10608-008-9202-1

Kessler, K., & Wang, H. (2012). Spatial perspective taking is an embodied process, but not for everyone in the same way: Differences predicted by sex and social skills score. *Spatial Cognition and Computation: An Interdisciplinary Journal*, *12*(2-3), 133–158. doi:10.1080/13875868.2011.634533

Kokkinara, E., & Slater, M. (2014). Measuring the effects through time of the influence of visuomotor and visuotactile synchronous stimulation on a virtual body ownership illusion. *Perception*, *43*(1), 43–58. doi:10.1068/p7545

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*. doi:10.3389/fpsyg.2013.00863

Lambrey, S., Doeller, C., Berthoz, A., & Burgess, N. (2012). Imagining being somewhere else: Neural basis of changing perspective in space. *Cerebral Cortex*, *22*, 166–174. doi:10.1093/cercor/bhr101

Langdon, R., Coltheart, M., Ward, P. B., & Catts, S. V. (2001). Visual and cognitive perspective-taking impairments in schizophrenia: A failure of allocentric simulation? *Cognitive Neuropsychiatry, 6*(4), 241–269. doi:10.1080/13546800143000005

Linacre, J. M. (n.d.). *Bonferroni - multiple t-tests*. Retrieved March 10, 2015, from http://www.winsteps.com/winman/bonferroni.htm

Longe, O., Maratos, F. A., Gilbert, P., Evans, G., Volker, F., Rockliff, H., & Rippon, G. (2010). Having a word with yourself: Neural correlates of self-criticism and self-reassurance. *NeuroImage, 49*(2), 1849–1856. doi:10.1016/j.neuroimage.2009.09.019

MacBeth, A., & Gumley, A. (2012). Exploring compassion: A meta-analysis of the association between self-compassion and psychopathology. *Clinical Psychology Review, 32*(6), 545–552. doi:10.1016/j.cpr.2012.06.003

Maselli, A., & Slater, M. (2013). The building blocks of the full body ownership illusion. *Frontiers in Human Neuroscience, 7*. doi:10.3389/fnhum.2013.00083

Mohr, C., Rowe, A. C., & Blanke, O. (2010). The influence of sex and empathy on putting oneself in the shoes of others. *British Journal of Psychology, 101*, 277–291. doi:10.1348/000712609X457450

Motulsky, H. (2011). *Do we need a global test before post hoc tests?* Retrieved February 24, 2015, from http://stats.stackexchange.com/questions/9751/do-we-need-a-global-test-before-post-hoc-tests

Neff, K. D. (2003). Self-Compassion: An alternative conceptualization of a healthy attitude toward oneself. *Self and Identity, 2*(2), 85–101. doi:10.1080/15298860309032

Neff, K. D., & Pommier, E. (2013). The relationship between self-compassion and other-focused concern among college undergraduates, community adults, and practicing meditators. *Self and Identity*, *12*(2), 160–176. doi:10.1080/15298868.2011.649546

Neff, K. D., Rude, S. S., & Kirkpatrick, K. L. (2007). An examination of self-compassion in relation to positive psychological functioning and personality traits. *Journal of Research in Personality*, *41*, 908–916. doi:10.1016/j.jrp.2006.08.002

Odou, N., & Vella-Brodrick, D. A. (2013). The efficacy of positive psychology interventions to increase well-being and the role of mental imagery ability. *Social Indicators Research, 110*, 111–129. doi:10.1007/s11205-011-9919-1

Opinio (version 6.8) [computer software]. (2014). ObjectPlanet. Retrieved from https://opinio.ucl.ac.uk/admin/folder.do

Pearson, A., Ropar, D., & Hamilton, A. F. (2013). A review of visual perspective taking in autism spectrum disorder. *Frontiers in Human Neuroscience, 7*. doi:10.3389/fnhum.2013.00652

Pearson, J., Rademaker, R. L., & Tong, F. (2011). Evaluating the mind's eye: The metacognition of visual imagery. *Psychological Science, 22*, 1535–1542. doi:10.1177/0956797611417134

Peck, T. C., Seinfeld, S., Aglioti, S. M., & Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and Cognition*, *22*(3), 779–787. doi:10.1016/j.concog.2013.04.016

Petkova, V. I., & Ehrsson, H. H. (2008). If I were you: Perceptual illusion of body swapping. *PLoS ONE, 3*(12). doi:10.1371/journal.pone.0003832

Pierce, J. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics, 2*. doi:10.3389/neuro.11.010.2008

Pierce, J. (2014). PsychoPy (version 1.80.03) [computer software]. Retrieved from http://www.psychopy.org

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *The Behavioral and Brain Sciences*, *4*, 515–526. doi:10.1017/S0140525X00076512

Raij, A., Kotranza, A., Lind, D. S., & Lok, B. (2009). Virtual experiences for social perspective-taking. In *IEEE Virtual Reality*. doi:10.1109/VR.2009.4811005

Ratcliff, G. (1979). Spatial thought, mental rotation and the right cerebral hemisphere. *Neuropsychologia*, *17*, 49–54. doi:10.1016/0028-3932(79)90021-6

Rector, N. A., Bagby, R. M., Segal, Z. V., Joffe, R. T., & Levitt, A. (2000). Self-criticism and dependency in depressed patients treated with cognitive therapy or pharmacotherapy. *Cognitive Therapy and Research*, *24*(5), 571–584. doi:10.1023/A:1005566112869

Rockliff, H., Gilbert, P., McEwan, K., Lightman, S., & Glover, D. (2008). A pilot exploration of heart rate variability and salivary cortisol responses to compassion-focused imagery. *Clinical Neuropsychiatry*, *5*(3), 132–139.

Rus-Calafell, M., Gutiérrez-Maldonado, J., Botella, C., & Baños, R. M. (2013). Virtual reality exposure and imaginal exposure in the treatment of fear of flying: A pilot study. *Behavior Modification*, *37*(4), 568–590. doi:10.1177/0145445513482969

Schulz, K. F., Altman, D. G., Moher, D., & Group, C. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomized trials.

Annals of Internal Medicine, *152*(11), 727–732. doi:10.7326/0003-4819-152-11-201006010-00232

Schurz, M., Aichhorn, M., Martin, A., & Perner, J. (2013). Common brain areas engaged in false belief reasoning and visual perspective taking: A meta-analysis of functional brain imaging studies. *Frontiers in Human Neuroscience, 7*. doi:10.3389/fnhum.2013.00712

Seltman, H. (2014). Mixed models: A flexible approach to correlated data. In *Experimental Design and Analysis* (pp. 357–378). Retrieved from http://www.stat.cmu.edu/~hseltman/309/Book/Book.pdf

Shahar, B., Szsepsenwol, O., Zilcha-Mano, S., Haim, N., Zamir, O., Levi-Yeshuvi, S., & Levit-Binnun, N. (2014). A wait-list randomized controlled trial of loving-kindness meditation programme for self-criticism. *Clinical Psychology and Psychotherapy*. doi:10.1002/cpp.1893

Singer, P. (2013). *Statistical significance tests on correlation coefficients*. Retrieved January 27, 2015, from http://www.philippsinger.info/?p=347

Slater, M., Spanlang, B., Sanchez-Vives, M. V., & Blanke, O. (2010). First person experience of body transfer in virtual reality. *PLoS ONE*, *5*(5), 1–9. doi:10.1371/journal.pone.0010564

Spagnolli, A., Bracken, C. C., & Orso, V. (2014). The role played by the concept of presence in validating the efficacy of a cybertherapy treatment: A literature review. *Virtual Reality*, *18*, 13–36. doi:10.1007/s10055-013-0241-x

Spreng, R. N., Mar, R. A., & Kim, A. S. N. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: A quantitative meta-analysis. *Journal of Cognitive Neuroscience*, *21*(3), 489–510. doi:10.1162/jocn.2008.21029

Srinivasan, N., & Gupta, R. (2011). Rapid communication: Global-local processing affects recognition of distractor emotional faces. *Quarterly Journal Of Experimental Psychology*, *64*(3), 425–433. doi:10.1080/17470218.2011.552981

Stellar, J. E., Manzo, V. M., Kraus, M. W., & Keltner, D. (2012). Class and compassion: Socioeconomic factors predict responses to suffering. *Emotion*, *12*, 449–459. doi:10.1037/a0026508

Strahan, R., & Gerbasi, K. C. (1972). Short, homogenous versions of the Marlow-Crowne Social Desirability Scale. *Journal of Clinical Psychology*, *28*(2), 191–193.

Surtees, A., Apperly, I., & Samson, D. (2013). Similarities and differences in visual and spatial perspective-taking processes. *Cognition*, *129*(2), 426–438. doi:10.1016/j.cognition.2013.06.008

Thakkar, K. N., & Park, S. (2010). Empathy, schizotypy, and visuospatial transformations. *Cognitive Neuropsychiatry*, *15*(5), 477–500. doi:10.1080/13546801003711350

Thirioux, B., Mercier, M. R., Blanke, O., & Berthoz, A. (2014). The cognitive and neural time course of empathy and sympathy: An electrical neuroimaging study on self-other interaction. *Neuroscience*, *267*, 286–306. doi:10.1016/j.neuroscience.2014.02.024

Thirioux, B., Tandonnet, L., Jaafari, N., & Berthoz, A. (2014). Disturbances of spontaneous empathic processing relate with the severity of the negative symptoms in patients with schizophrenia: A behavioural pilot-study using virtual reality technology. *Brain and Cognition*, *90*, 87–99. doi:10.1016/j.bandc.2014.06.006

Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences*, *16*, 495–552. doi:10.1017/S0140525X0003123X

Tversky, B., & Hard, B. M. (2009). Embodied and disembodied cognition: Spatial perspective-taking. *Cognition*, *110*(1), 124–129. doi:10.1016/j.cognition.2008.10.008

Vastano, R., Sulpizio, V., Steinisch, M., Comani, S., & Committeri, G. (2014). Embodied and disembodied allocentric simulation in high schizotypal subjects. *Experimental Brain Research*, *232*, 3023–3033. doi:10.1007/s00221-014-3991-0

Wallach, H. S., Safir, M. P., & Bar-Zvi, M. (2009). Virtual reality cognitive behavior therapy for public speaking anxiety: A randomized clinical trial. *Behavior Modification*, *33*(3), 314–338. doi:10.1177/0145445509331926

Won, A. S., Bailenson, J., Lee, J., & Lanier, J. (2015). Homuncular flexibility in virtual reality. *Journal of Computer-Mediated Communication*. doi:10.1111/jcc4.12107

Zhang, X., Yu, H. W., & Barrett, L. F. (2014). How does this make you feel? A comparison of four affect induction procedures. *Frontiers in Psychology*, *5*. doi:10.3389/fpsyg.2014.00689

Zou, G. Y. (2007). Toward using confidence intervals to compare correlations. *Psychological Methods*, *12*(4), 399–413. doi:10.1037/1082-989X.12.4.399

Part 3: Critical Appraisal

1: Introduction

This appraisal reflects on the process of carrying out part 1, the literature review and part 2, the empirical study. First, the literature and processes involved in conducting the meta-analysis in part 1 are discussed. Second, the following methodological choices in part 2 are expanded on: removing a condition, sample size, design of the mental imagery (MI) intervention, and the measures used. Finally, issues with using immersive virtual reality (IVR) in the empirical study and in psychological therapy more generally are considered.

2: The Literature Review

2.1 The Literature

Throughout my clinical psychology training I have been interested in mechanisms of therapeutic change and dismantling research methodology. Having completed the proposal for the empirical study I was aware that the use of IVR as a therapeutic tool was a rapidly growing, if experimental, field of psychotherapy research (Riva, 2005). So when tasked with reviewing an area of literature related to the empirical study, I decided that it would be relevant and interesting to review investigations of mechanisms underlying the effectiveness of psychological treatment using IVR. Scoping searches revealed that the sense of presence in IVR was frequently investigated as a possible mechanism and findings for its role in the efficacy of treatment using IVR were mixed. Therefore, this seemed a fitting focus for the review. It became apparent that different research groups had developed different theories of presence. Additionally, they had then developed a self-report

measure which tapped their particular theoretical construct. Furthermore, presence in IVR had been investigated across a range of settings and treatments using IVR using a variety of methodologies. It seemed challenging to draw together this heterogeneous literature. However, it also seemed that systematic examination of the literature could elucidate differences in findings and highlight issues with its interpretation.

2.2 Conducting the Meta-Analysis

Given variation in the studies' findings a meta-analytic approach seemed appropriate. Statistical analysis allowed estimate of the effect size across studies and formal assessment of the consistency of the findings from one type of study to the next. This approach also provided a transparent, objective and replicable framework for synthesis of the data (Borenstein et al., 2009). However, initial readings of the studies revealed variations in their design and methodological quality which could not be captured by meta-analysis alone. It was considered important to also assess these factors to aid interpretation of any findings.

Attempts were made to contact corresponding authors for studies which had missing data. Only one author replied and supplied data. She then referred me to her colleague who provided further information. Lack of response from other authors meant that certain studies could not be included in the analysis. At this stage, I considered whether an entirely narrative approach to the review would be preferable. My initial rationale for taking a statistical approach remained, but it seemed important to include the results of studies with incomplete data to provide a comprehensive account of the literature. As such, I decided to conduct a meta-

analysis accompanied by a summary of the results of studies which were ineligible for inclusion.

Some studies reported multiple treatment outcomes for the same sample. Meta-analysis assumes independence of effects but different treatment outcomes for the same participants are not independent (Borenstein et al., 2009). This issue was dealt with by generating a single, mean treatment outcome statistic for each sample for entry into the meta-analysis. With hindsight, it was noted that an average correlation that comes from several very different correlations may not have the same contribution to a meta-correlation as an average correlation that comes from several similar correlations. Published meta-analyses address this unit of analysis issue in different ways. Gentes and Ruscio (2011) used the above approach and did not mention this concern. Ling, Nefs, Morina, Heynderickx, and Brinkman (2014), who investigated anxiety disorders, selected the anxiety measure with the highest anxiety score in cases where a sample had completed multiple measures of anxiety. Another approach is to use the effect size and variance for each outcome to create a synthetic (average) effect size (Borenstein et al., 2009). The approach taken for the meta-analysis was considered the most feasible solution of those outlined above given that there was no clear rationale for selecting scores from one outcome measure over another and computing the variance involved highly complex calculations which were beyond the reach of the review.

Two studies ran regression analyses which required transformation in order to be included. This process could also have been carried out in different ways. There appeared to be no consensus on which transformation was best as different researchers argued for the merits of different methods (Aloe & Becker, 2011; Peterson & Brown, 2005). Transformation to a semi-partial correlation was

ultimately considered the most coherent and established method. However, this statistic contains the influence of the other predictors in the regression analysis. Therefore, it was decided to run two meta-analyses, one including and one excluding the transformed data, to provide an informative and transparent statistical evaluation.

The range of approaches used to address both of the above statistical issues reflects the reality of using developing statistical methodologies. A choice must be made between ways to address the issue, none of which offer a perfect solution. Making these decisions was challenging as it required researching and evaluating possible statistical solutions rather than following an established approach to analysis. The review was initially based on a relatively small number of studies. Some decisions led to a reduction in the number of studies which were included in particular aspects of it. I had not anticipated this at the outset. Whilst a small number of studies can affect the validity of the findings, the choices I made seemed appropriate in terms of producing a comprehensive review of the literature.

3: The Empirical Study

Some aspects of the empirical study were born out of recent areas of research such as the use of IVR to cultivate self-compassion. Other aspects were novel such as investigation of the link between visuo-spatial perspective-taking (VSPT) and the ability to cultivate self-compassion using IVR or MI. The exploratory nature of the study raised interesting methodological questions and challenges which are discussed below.

3.1 Removing a Condition

The proposal for the study included a third condition, a pre-existing non-immersive virtual reality intervention (nIVR). It was planned that participants would be randomly assigned to each of the three conditions (IVR, MI, nIVR) in a 1:1:1 ratio, stratified by gender. The nIVR intervention would have taken place in the same part of the lab as the other conditions and required participants to sit in front of a TV screen with headphones on. The intervention would have involved: delivering the three step compassionate response to a 2D image of a child avatar on the TV screen (the same child avatar used in the IVR condition); observing the child's response (the same response as in the IVR condition); receiving compassion by experiencing a recorded replay of the compassionate response, hearing one's voice and seeing one's face in 2D on the TV screen. This condition was excluded before data collection began. Although it was intended to be an analogue of the other interventions it was decided that it introduced confounds such as seeing one's own face. Also, there seemed no clear rationale for it to meaningfully contribute to the study aims.

3.2 Sample Size

Sample size was originally computed prior to the decision to drop the nIVR condition. This calculation used the same parameters reported in the empirical paper except the number of groups was three (between subjects: IVR, MI, nIVR). This generated an overall sample size of 54 (18 per condition). It was decided to test a sample size of 20 per condition to allow for errors in data collection. After the nIVR condition was removed the need to re-calculate sample size was overlooked and not recognised until after the data collection phase had ended. At this stage a re-

calculation was carried out to check the power of the study. This calculation (reported in the empirical paper) generated a sample size of 42, meaning that each condition in the study was underpowered by 1 participant. Through carrying out this research I have gained a greater understanding of the processes involved including the implications of methodological decisions. If undertaking such research again I would re-run the power calculation straight after a methodological decision which involved change to the design of the study.

3.3 Designing the Mental Imagery Condition

The IVR scenario was a pre-existing intervention. The scenario was designed in consultation with clinical psychologists that are experts in the field of compassion and was intended to be accessible without prior therapeutic input (Falconer et al., 2014). Much consideration was given to the design of the MI condition. Use of a Compassion Focused Therapy (CFT) MI exercise intended for use in clinical practice was considered. A review of the literature found that CFT MI work designed to cultivate self-compassion typically started by inviting the individual to focus on compassionate qualities and allow an image of these qualities to come to mind. After development and practice of these images, MI exercises often progressed to generating images of a compassionate part of oneself or, if preferred, to experience the image *as if* an external other is comforting you (Gilbert & Irons, 2005). For example, one brief CFT exercise designed to cultivate a compassionate-self guided the individual to create an image of themselves at their compassionate best imbued with the specific qualities of compassion. Next it asked the individual to imagine looking at this compassionate-self from the outside; seeing their behaviour, noticing their motivations and noticing others' responses to them (Gilbert & Choden, 2013).

CFT MI work took place over several sessions and even the above brief scenario was proceeded by practices which included mindfulness training, the cultivation of positive emotional systems, and developing compassionate images. Exploration of the literature raised the following questions around the suitability of using a CFT MI exercise: Would its use as a one-off intervention be accessible and meaningful without the prior stages used in CFT training? If the scenario in each condition was different how valid was their comparison? How would this effect the interpretation of any findings?

The alternative was to develop an MI analogue of the IVR scenario. This distanced the MI scenario from clinical practice. However, it facilitated comparison with the IVR scenario and seemed to improve the feasibility of cultivating compassion without prior CFT training. This approach also appealed to concerns about interpreting findings. Therefore, it was decided to develop an MI analogue of the IVR intervention. However, the IVR scenario included guided exercises designed to facilitate embodiment of the avatar and enable participants to become accustomed with the virtual environment. So, in order to achieve as much correspondence with the IVR scenario as possible, these exercises were included in the MI scenario. This somewhat superfluous consequence of developing analogue scenarios may have impacted on the efficacy of the MI condition. However, with hindsight, the decision to use analogue scenarios seems appropriate in terms of statistical and practical considerations given the overall purpose of the study.

3.4 Measures

The Four Mountains tasks and Manikin task were originally in a format that required manual administration. For standardisation, removal of researcher bias,

ease of administration, accurate data collection and to increase the ease of data analysis I decided to use programming software to convert these tasks into a computerised format. This enabled the tasks to be completed in the absence of a researcher. This was a time-consuming but useful exercise. It fulfilled the rationale for computerised administration and during the programming process I began to reflect on the VSPT requirements of the Four Mountains task (discussed below). Disadvantages of this method of administration were that differences between participants' level of engagement were not observed and their engagement may have differed had a researcher been present.

The Four Mountains allocentric VSPT task was chosen as it was freely available and both easy and practical to administer in the university virtual reality lab. Firstly, this task required allocentric processing of global landscape features. Srinivasan and Gupta (2011) investigated the effect of global–local processing on the recognition of faces with happy or sad emotional expressions. They demonstrated that an experimental task of perceptual processing without emotional content was associated with the processing of emotional face stimuli. This supports the validity of relating an experimental VSPT task to emotional stimuli in the empirical study. However, their results showed that local processing facilitated recognition of sad faces. The intervention in the empirical study involved giving compassion to a crying child and then taking their perspective. Therefore, an allocentric VSPT task requiring localised processing of small scale cues such as the Virtual Town Square test (King, Burgess, Hartley, Vargha-Khadem, & O'Keefe, 2002) may have been an interesting addition to the assessment battery.

Secondly, experimental evidence shows that the presence of another person in a visual scene elicits an allocentric remapping of space with reference to the other

person, even in the absence of explicit instruction to adopt their perspective (Becchio et al., 2013; Tversky & Hard, 2009). Interestingly, Vastano, Sulpizio, Steinisch, Comani, and Committeri (2014) found evidence of a difficulty with disembodied but not embodied allocentric VSPT in healthy individuals with high levels of schizotypal personality traits. The authors suggest that the absence of a deficit in embodied processing may have been due to the facilitating effect of including the person, whose perspective is to be adopted, in the visual scene. The scenario in the empirical study involved VSPT in the presence of another person in the visual scene. The above findings suggest that this may have stimulated and indeed facilitated embodied allocentric VSPT. However the Four Mountains task assessed disembodied allocentric VSPT, a subtly different ability. These reflections suggest that the specific type of allocentric processing required by the task may have differed from that required by the intervention and therefore limited findings. Whilst measures included in the empirical study were constrained by the factors mentioned above, it seems that the study may have benefited from a more detailed consideration of how best to assess allocentric VSPT ability.

The Virtual Reality Experience Questionnaire (VREQ) was chosen to measure the sense of embodiment as it was designed by Falconer et al. (2014) for their study which used the same IVR scenario. Kilteni, Groten, and Slater (2012) propose that one experiences a sense of embodiment if at least one of the following three senses are experienced with minimal intensity: self location inside the virtual body, one feels to be an agent of the virtual body, one feels the virtual body is one's own body (body ownership). The VREQ measures all aspects of embodiment. Analysing all three aspects fitted with the ethos of the empirical study which was to comprehensively explore the factors being investigated. However, it seemed that the

amount of correlational analysis this required would have considerably hindered meaningful interpretation. Therefore it was decided to measure one aspect of the sense of embodiment. The illusion of body ownership was chosen as this aspect is concerned with the engagement of processes that make us feel we own our biological bodies. This adds to the reality of what is being perceived (Kokkinara & Slater, 2014; Maselli & Slater, 2013). It would be interesting to extend the empirical study by investigating all three aspects. However it is worth noting that all aspects were analysed by Falconer et al. (2014) and none were found to be responsible for change in state self-compassion or state self-criticism following the IVR intervention. Replication would confirm this finding. It may be also be worthwhile developing alternative measures in order to fully explore the contribution of the sense of embodiment to IVR interventions designed to cultivate self-compassion.

The use of self-report measures raised the issue of social desirability bias. The existing literature highlighted this concern in relation to some of the concepts investigated in the empirical study. For example, Pearson, Rademaker, and Tong (2011) found that individuals could reliably evaluate the vividness of single episodes of imagination using an experimental paradigm and a revised version of the Vividness of Visual Imagery Questionnaire (Marks, 1973), a self-report measure of image vividness. However, Allbutt, Ling, Rowley, and Shafiullah (2011) found that a version of this measure correlated significantly with measures of socially desirable responding. Whilst these concerns were considered, development and use of alternative ways of assessing such constructs was beyond the scope of the study.

4: Immersive Virtual Reality

The current study indicated that state self-compassion increased following the MI intervention. The IVR system used in the empirical study suffered from several limitations. Technological limitations included the avatars having a set facial expression and when embodied in the avatar and looking down at the floor occasionally it was possible to see through the avatar's body. When embodied in the adult avatar and giving compassion to the child, participants often moved, sometimes getting closer to the child, bending down to the child's level or reaching out to the child. These movements were regarded as important expressions of compassion to be re-experienced. Therefore, as the child avatar participants experienced a recorded replay of their adult avatar. Although participants were asked to stand in a particular location in the room when they embodied the child avatar, at times their movement as the adult avatar led the adult to appear at an odd angle or distance to the child. It was observed that some participants spent time trying to reposition themselves in relation to the adult during the recorded replay. These factors, in addition to those discussed in the empirical paper, may have impacted on the ability of the IVR intervention to cultivate compassion. They also acted as extraneous factors which may have affected the comparability of the IVR and MI conditions.

Development of the IVR scenario may offer some benefits. By embodying the avatars individuals who otherwise would resist imagining feelings of self-compassion may be able to access this experience. More generally, IVR has the potential to be used by individuals across the lifespan. It may be an appealing therapeutic tool for some individuals which could therefore improve the rates at which people seek treatment and treatment compliance (Price, Anderson, &

Rothbaum, 2008). Technological developments may further increase the ability to generate realistic virtual environments which may enhance ecological validity and generalisation to real-world situations. The scope for flexible, interactive and individualised scenarios and the convenience and control offered by IVR further add to the potential value of this tool for psychological treatment (Gregg & Tarrier, 2007).

Despite this, current research suggests a need for caution in carrying out treatment using IVR. Interestingly, some aspects of the IVR intervention were experienced negatively and it was observed that some participants appeared slightly anxious about experiencing the IVR. This could indicate that for some individuals immersion in IVR may be aversive. This is a tentative suggestion that requires investigation. However, it highlights that in addition to selecting the most efficacious treatment tool based on existing research, careful consideration should be given to an individual's suitability for treatment using a particular tool. Whilst preparation for undertaking this treatment may be beneficial it seems that some of the present limitations of IVR technology may negatively impact on its ability to offer therapeutic benefit. Using this technology presents other challenges such as affordability, space for and maintenance of the equipment and training in its use. These concerns along with the results of the literature review and the empirical study suggest that much research and development of IVR as a therapeutic tool is needed before it can be considered for use in clinical practice.

5: Conclusion

The process of carrying out the research raised many challenges. These included considering how best to draw together and draw from the existing literature, how best to provide a platform for further investigation, and how the research could ultimately inform clinical practice. Overall, I think the literature review and empirical study addressed their aims. Conducting the research was a rewarding experience that furthered my understanding of research processes and developed my ability to critically consider and interpret the evidence-base for clinical practice.

References

Allbutt, J., Ling, J., Rowley, M., & Shafiullah, M. (2011). Vividness of visual imagery and social desirable responding: Correlations of the vividness of visual imagery questionnaire with the balanced inventory of desirable responding and the Marlowe-Crowne scale. *Behavior Research Methods, 43*, 791–799. doi:10.3758/s13428-011-0086-8

Aloe, A. M., & Becker, B. J. (2011). An effect size for regression predictors in meta-analysis. *Journal of Educational and Behavioral Statistics, 37*(2), 278–297. doi:10.3102/1076998610396901

Becchio, C., Del Giudice, M., Dal Monte, O., Latini-Corazzini, L., & Pia, L. (2013). In your place: Neuropsychological evidence for altercentric remapping in embodied perspective taking. *Social Cognitive and Affective Neuroscience, 8*, 165–170. doi:10.1093/scan/nsr083

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: John Wiley & Sons.

Falconer, C. J., Slater, M., Rovira, A., King, J. A., Gilbert, P., Antley, A., & Brewin, C. R. (2014). Embodying compassion: A virtual reality paradigm for overcoming excessive self-criticism. *PLoS ONE, 9*(11). doi:e111933. doi:10.1371/journal.pone.0111933

Gentes, E. L., & Ruscio, A. M. (2011). A meta-analysis of the relation of intolerance of uncertainty to symptoms of generalized anxiety disorder, major depressive disorder, and obsessive-compulsive disorder. *Clinical Psychology Review, 31*(6), 923–933. doi:10.1016/j.cpr.2011.05.001

Gilbert, P., & Choden. (2013). *Mindful compassion*. London: Constable & Robinson Ltd.

Gilbert, P., & Irons, C. (2005). Focused therapies and compassionate mind training for shame and self-attacking. In P. Gilbert (Ed.), *Compassion: Conceptualisations, research and use in psychotherapy* (pp. 263–325). East Sussex: Routledge.

Gregg, L., & Tarrier, N. (2007). Virtual reality in mental health: A review of the literature. *Social Psychiatry and Psychiatric Epidemiology*, *42*, 343–354. doi:10.1007/s00127-007-0173-4

Kilteni, K., Groten, R., & Slater, M. (2012). The sense of embodiment in virtual reality. *Presence, 21*(4), 373–387. doi:10.1162/PRES_a_00124

King, J. A., Burgess, N., Hartley, T., Vargha-Khadem, F., & O'Keefe, J. (2002). Human hippocampus and viewpoint dependence in spatial memory. *Hippocampus*, *12*, 811–820. doi:10.1002/hipo.10070

Kokkinara, E., & Slater, M. (2014). Measuring the effects through time of the influence of visuomotor and visuotactile synchronous stimulation on a virtual body ownership illusion. *Perception, 43*(1), 43–58. doi:10.1068/p7545

Ling, Y., Nefs, H. T., Morina, N., Heynderickx, I., & Brinkman, W. P. (2014). A meta-analysis on the relationship between self-reported presence and anxiety in virtual reality exposure therapy for anxiety disorders. *PloS ONE*, *9*(5). doi:10.1371/journal.pone.0096144

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, *64*(1), 17–24.

Maselli, A., & Slater, M. (2013). The building blocks of the full body ownership illusion. *Frontiers in Human Neuroscience, 7*. doi:10.3389/fnhum.2013.00083

Pearson, J., Rademaker, R. L., & Tong, F. (2011). Evaluating the mind's eye: The metacognition of visual imagery. *Psychological Science, 22*, 1535–1542. doi:10.1177/0956797611417134

Peterson, R. A., & Brown, S. P. (2005). On the use of beta coefficients in meta-analysis. *The Journal of Applied Psychology*, *90*(1), 175–181. doi:10.1037/0021-9010.90.1.175

Price, M., Anderson, P., & Rothbaum, B. O. (2008). Virtual reality as treatment for fear of flying: A review of recent research. *International Journal of Behavioral Consultation and Therapy*, *4*(4), 340–347. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ861345&site=ehost-live

Riva, G. (2005). Virtual reality in psychotherapy: Review. *CyberPsychology & Behavior, 8*(3), 220–230. doi:10.1089/cpb.2005.8.220

Srinivasan, N., & Gupta, R. (2011). Rapid communication: Global-local processing affects recognition of distractor emotional faces. *Quarterly Journal Of Experimental Psychology, 64*(3), 425–433. doi:10.1080/17470218.2011.552981

Tversky, B., & Hard, B. M. (2009). Embodied and disembodied cognition: Spatial perspective-taking. *Cognition, 110*(1), 124–129. doi:10.1016/j.cognition.2008.10.008

Vastano, R., Sulpizio, V., Steinisch, M., Comani, S., & Committeri, G. (2014). Embodied and disembodied allocentric simulation in high schizotypal subjects. *Experimental Brain Research, 232*, 3023–3033. doi:10.1007/s00221-014-3991-0

Appendices

Appendix A: Studies Excluded From the Literature Review

Table A1

Characteristics of the Excluded Studies

Study	**Reason for exclusion**
Emmelkamp, Bruynzeel, Drost, and van der Mast (2001)	No analysis of the relationship between presence and the outcome of treatment using IVR
Gold, Kim, Kant, Joseph, and Rizzo (2006)	No analysis of the relationship between presence and the outcome of treatment using IVR
Grewe et al. (2013)	Did not investigate a treatment using IVR
Hodges et al. (1995)	No analysis of the relationship between presence and the outcome of treatment using IVR
Hoffman, Garcia-Palacios, Carlin, Furness, and Botella-Arbona (2003)	No analysis of the relationship between presence and the outcome of treatment using IVR
Krijn et al. (2004)	No analysis of the relationship between presence and the outcome of treatment using IVR
Krijn et al. (2007)	No analysis of the relationship between presence and the outcome of treatment using IVR
Robillard, Bouchard, Fournier, and Renaud (2003)	No analysis of the relationship between presence and the outcome of treatment using IVR
Sharar et al. (2007)	No analysis of the relationship between presence and the outcome of treatment using IVR
Villani, Riva, and Riva (2007)	Healthy sample
Wiederhold, Davis, and Wiederhold (1998)	No analysis of the relationship between presence and the outcome of treatment using IVR

Appendix B: The Author's Contribution to the Joint Project

The empirical study presented in this thesis was part of a joint project conducted by myself (Nicola Alden, author of this thesis) and a second Trainee Clinical Psychologist, Anneka Holden. We designed the project together and jointly carried out the following preparatory tasks: writing the ethics amendment document, writing participant information sheets, writing and recording the guided mental imagery condition and familiarisation with the immersive virtual reality technology. Anneka took a lead role in setting up the online questionnaires in Opinio, Version 6.8, whilst I took a larger role in recruiting participants. Anneka identified participants who were eligible to take part in the experimental session based on data from the screening survey. I emailed all participants to inform them whether or not they had been selected to take part in the experimental session and scheduled the sessions. The majority of the experimental sessions were run jointly. At the end of each session I took care of the technical equipment and other resources, whilst Anneka set up text message reminders for the participants. Data analysis and writing-up of all parts of the thesis was carried out completely independently.

Appendix C: Participant Information Sheets

Volunteer Information Sheet

You will be given a copy of this information sheet.

Title of project: Approaches to Nurturing Compassion

This study has been approved by the UCL Research Ethics Committee. [Project ID Number: DSD.2013.010]

Purpose of the study:
The purpose of this study is to investigate whether virtual reality (VR) and mental imagery can be used to influence the experience of compassion. It will also assess whether visual perspective-taking ability is related to this process.

Investigators:
Prof Chris Brewin, Dr John King, Anneka Holden, Nicola Alden

We would like to invite you to participate in this research project. You should only participate if you want to; choosing not to take part will not disadvantage you in any way. Before you decide whether you want to take part, it is important for you to read the following information carefully. Ask us if there is anything that is not clear or if you would like more information.

----------Please read the following carefully ----------

Background of the study:
Compassion is an inherent human emotion. New research shows that compassion plays an important role in our lives and can influence our general well-being. This research has also highlighted that we can be both compassionate to others and also to ourselves. For some people delivering compassion to the self or others can be difficult or awkward. For other people being compassionate is relatively easy. We are interested in helping people become more compassionate, especially as it has been shown to positively impact our psychological health.

The aim of this study is to investigate whether VR and mental imagery can be used to influence the experience of compassion. It will also assess whether visual perspective-taking ability is related to this process.

Who can participate?
We are looking for healthy males and females, with no history of mental illness or brain damage, between the ages of 18 – 50. We will be selecting individuals who have average to above average self-criticism levels.

Do I have to take part?
It is up to you to decide whether you wish to take part. Please know that your information is kept in confidence, that your data will not be personally identifiable, and that you are free to withdraw at any time, without giving a reason.

What will happen to me if I decide to take part?
If you decide to take part, you will be required to sign a consent form and fill out several questionnaires about self-attitudes and your emotions. This can be completed at home on your computer and will take you approximately 20 minutes. After this you will be invited to take part in either a VR or a mental imagery experiment session at UCL. There you will have to answer four short questionnaires and complete two perspective-taking tasks. After this you will be randomly allocated to one of the following tasks:

1. **An immersive VR experiment:** This involves putting on a lightweight suit and a head mounted display. Through this equipment you will see a virtual world in which you will see an "avatar" (a movable three-dimensional image that represents a person in a virtual reality environment) of yourself and a child. Your task will be to interact compassionately with the child avatar by talking to him/her. We will provide you with instructions on how to go about this. When you have done this you will then re-experience your compassionate interaction from the child's perspective.

2. **A mental imagery experiment:** Your task will be the same as above except that you will hear an audio recording which will guide you to imagine interacting with a child.

The experimental session will take 40-60 minutes.

During the VR experiments we will record your verbal responses. At the end of the session we will provide you with debriefing information and you will be entered into a prize draw for Amazon vouchers. There are 19 prizes ranging from £100 to £10 (we will give you further details on the day). If you are a UCL undergraduate student you will also receive course credits for your participation.

After you have completed the experimental task you will be asked to fill in five more short questionnaires. You will also be asked to practice imagining the experimental task regularly for two weeks following the session. You will be sent an automated text message every other day reminding you to do this. At the end of the two weeks you will be requested to complete several questionnaires at home on your computer.

What are the possible disadvantages and risks of taking part?
The task and the questionnaires used in this study are regarded as innocuous for healthy participants. However, if at any stage you wish to stop the experiment then you may do so. We will also have a clinical psychologist (Prof Chris Brewin) on hand should you feel the need to talk to someone.

People can sometimes experience a degree of nausea when using virtual reality. If you feel nauseous please say so and we can stop the experiment.

There has been some research showing that the use of head mounted displays can disturb vision – up to approximately 30 minutes after use. This risk is small and no long term effects would be expected. However, we would ask that you take precaution after the experiment.

There have also been reports that virtual reality can induce flashbacks and epileptic seizures in vulnerable individuals. If you feel like you might be at a particular risk to either of these we would ask you not to participate.

What are the possible benefits of taking part?
You will have the opportunity to experience, first hand, cutting-edge technology used to deliver virtual reality. You will also contribute to the development of novel psychological treatments.

Will my taking part in the study be kept confidential?
Yes. Your information will be completely confidential. You will be assigned a unique participant number so that your data will not be personally identifiable. We will also follow ethical and legal practice and all information about you will be handled in confidence. All data will be collected and stored in accordance with the Data Protection Act 1998. This means that only the investigators will have access to the data from the study.

What if there is a problem?
If you have a concern about any aspect of this study, you should ask to speak to:

Anneka Holden and Nicola Alden (Trainee Clinical Psychologists)
Email: anneka.holden@ucl.ac.uk or n.alden.12@ucl.ac.uk
or
Prof. Chris R. Brewin (Clinical Psychologist)
Email: c.brewin@ucl.ac.uk

Clinical, Educational & Health Psychology
University College London
Gower Street, WC1E 6BT
London, U.K.

Information about the Session

You are about to take part in a virtual reality (VR) experiment. Before you start it is essential that you read the information below carefully. If you have any questions please don't hesitate to ask the researchers.

In general
The VR experiment involves putting on a lightweight suit and a head mounted display. Through this equipment you will see a virtual world in which you will see an "avatar" (a movable three-dimensional image that represents a person in a virtual reality environment) of yourself and a child. Your task will be to interact compassionately with a child avatar by talking to him/her. We will provide you with instructions on how to go about this. When you have done this you will then re-experience your compassionate interaction from the child's perspective.

The process
First of all you will be asked to complete some questionnaires and tasks about thoughts, beliefs and aptitudes that you may have.

Next, you will read some instructions about how to interact with the child avatar compassionately.

Then you will put on the suit. We will attach light reflecting balls to it, which will allow us to track your body position in the suit. This will only take a minute. The suit needs to be tight but it is stretchy. If you feel that it is very uncomfortable please tell us. You can keep your clothes on underneath or remove items of clothing if this will make you feel more comfortable.

Once you have put the suit on the researchers will calibrate the VR system. This will take a few moments and requires you to stand and walk about the room. We will talk you through this.

You will then be given some time to re-read the instructions. After this you will complete a five minute guided relaxation exercise. Then you will hear an audio recording asking you to carry out a few specific movements to help you familiarise yourself with the virtual environment and your avatar through the head mounted display. Afterwards we will give you a short amount of time to continue doing this. You can walk around a little, move your limbs and look in the mirror.

The VR session consists of three stages. The first is delivering your three stage compassionate response to the child avatar. The second stage involves a change of perspective, from your perspective to the child's perspective. You will then hear the same audio recording as before and be given a short amount of time to familiarise yourself with the environment from this new perspective. In the third stage you will experience your compassionate responses from the new perspective.

After the session you will be asked to complete a few more questionnaires.

When you are ready to continue please inform your administrator.

Instructions

Once you are ready the visual scene in the head mounted display will fade out and back in again. Within the new environment you will now be standing across from a child who is upset and crying. We would like you to interact compassionately with the child by comforting and talking to her/him.

Although this seems like a simple task many people have never been taught how to give compassion and may initially feel a little awkward in this situation. Research suggests that when trying to comfort someone in this way there are three essential steps. We would like you to use this three step procedure. Take a few moments now to understand and remember these three steps, and feel free to talk to your researcher about them:

The first stage is **validation**. The aim is to acknowledge that the other person is upset, that you do not judge them for this, and that it is perfectly acceptable for them to react in this way.

The second stage is **redirection of attention**. The aim is to direct the other person's attention towards something that is positive, soothing, and comforting.

The third stage is **memory activation**. The aim is to suggest that the person could try to recall a memory of someone who love them or is kind to them. This memory is supposed to instil positive feelings of warmth, comfort, and safety.

On the next page are several sentences that you can use when comforting the child. When talking to the child we would like you to talk **slowly, softly, and compassionately**. It is important that you **try not to rush your sentences**. It is also important to stay engaged with the person you are being compassionate towards: remain attentive to the child to convey that you are fully aware of their distress. We understand that this might be difficult or awkward for some people but please try your best.

After delivering a stage of the compassionate response we would like you to take a few moments to allow the child to absorb what you have said. In addition to this we would like you to observe the child for any changes in her/his behaviour in response to what you have said. For example, research shows that when recovering from being upset, people are likely to cease crying, remove their hands away from their

faces, lift their head up and then finally have a more upright posture and make eye contact with you when they are fully comforted. However, people respond to compassion in different ways and for some this may be a slow process. Keep in mind that it is compassionate to let people respond at their own pace. When you feel like the child has had enough time to absorb and respond to what you have said please proceed to the next step of your compassionate response.

Three-Step Compassionate Response

1. **Validation**
"It's not nice when things happen to us that we don't like. It's really upset you hasn't it?"

2. **Redirection of Attention**
"Sometimes when we are sad it's helpful to think of someone who loves us or is kind to us."

3. **Memory Activation**
"Can you think of someone who loves you or is kind to you? What might they say to you now that would make you feel better?"

Please take a few moments to remember these sentences as you will use them in the experiment. You do not need to remember them word for word, an approximate version is fine but please try to follow the script as closely as possible. Once you feel confident that you can deliver the sentences in a compassionate way please tell the researchers so that you can practice them together.

Information about the Session

You are about to take part in a mental imagery (MI) experiment. Before you start it is essential that you read the information below carefully. If you have any questions please don't hesitate to ask the researchers.

In general
Your task is to imagine interacting compassionately with a child, by talking to her/him in your head. When you have done this you will be asked to imagine experiencing your compassionate interaction from the child's perspective. We will provide you with instructions on how to go about this. During the experiment you will be seated, with your eyes closed and wearing headphones. You will hear audio instructions to guide you in imagining the scenario.

The process
First of all you will be asked to complete some questionnaires and tasks that will assess some thoughts, beliefs and aptitudes that you may have.

Next, you will read some instructions about how to interact compassionately with the child that you will imagine. After this you will complete a five minute guided relaxation exercise. You will then hear an audio recording asking you to carry out a few imagery tasks to help you familiarise yourself with the experiment.

The MI session consists of three stages which you will be guided through this by an audio recording. The first is delivering your three stage compassionate response to the child. The second stage involves imagining a change of perspective, from your perspective to the child's perspective. You will then be asked to complete similar imagery tasks as before to familiarise yourself with the third stage of the experiment. In this final stage you will imagine experiencing your compassionate responses from the child's perspective.

After the session you will be asked to complete a few more questionnaires.

When you are ready to continue please inform your administrator.

Instructions

Once you are seated on the stool with the headphones on you will be asked to imagine a child seated opposite you and then to interact with and comfort the child by talking to her/him in your head.

Although this seems like a simple task many people have never been taught how to give compassion and may initially feel a little awkward in this situation. Research suggests that when trying to comfort someone in this way there are three essential steps. We would like you to use this three step procedure. Take a few moments now to understand and remember these three steps, and feel free to talk to your researcher about them:

The first stage is **validation**. The aim is to acknowledge that the other person is upset, that you do not judge them for this, and that it is perfectly acceptable for them to react in this way.

The second stage is **redirection of attention**. The aim is to direct the other person's attention towards something that is positive, soothing, and comforting.

The third stage is **memory activation**. The aim is to suggest that the person could try to recall a memory of someone who love them or is kind to them. This memory is supposed to instil positive feelings of warmth, comfort, and safety.

On the next page are several sentences that you can say in your head to comfort the child. When talking to the child we would like you to **talk slowly, softly, and compassionately**. It is important that you **try not to rush your sentences**. It is also important to stay engaged with the person you are being compassionate towards: remain attentive to the child to convey that you are fully aware of their distress. We understand that this might be difficult or awkward for some people but please try your best.

After delivering a stage of the compassionate response you will be instructed to take a few moments to imagine that the child is absorbing what you have said. In addition to this you will be informed that the child has changed her/his behaviour in response to what you have said and you will be asked to imagine this. Research shows that when recovering from being upset, people are likely to cease crying, remove their hands away from their faces, lift their head up and then finally sit upright and make eye contact with you when they are fully comforted. However, people respond to compassion in different ways and for some this may be a slow process. Take some time to imagine that the child has absorbed and responded to what you have said and then proceed to the next step of your compassionate response.

In the second part of the task you will be asked to imagine that you are the child and you can see your adult-self seated opposite you. Then you will be asked to imagine that your adult-self is saying the three stage response to you in the same way as you said it to the child before.

Three-Step Compassionate Response

1. **Validation**
"It's not nice when things happen to us that we don't like. It's really upset you hasn't it?"

2. Redirection of Attention
"Sometimes when we are sad it's helpful to think of someone who loves us or is kind to us."

3. Memory Activation
"Can you think of someone who loves you or is kind to you? What might they say to you now that would make you feel better?"

Please take a few moments to remember these sentences as you will use them in the experiment. You do not need to remember them word for word, an approximate version is fine but please try to follow the script as closely as possible. Once you feel confident that you can deliver the sentences in a compassionate way please tell the researchers so that you can practice them together.

Participant Debrief Sheet

Title of project: Approaches to Nurturing Compassion

Compassion can be defined as concern for the suffering and misfortunes of others and is generally associated with caring, warmth and sympathy for others. Self-compassion is "the ability to hold one's [own] feelings of suffering with a sense of warmth, connection and concern" (Neff & McGehee, 2010). Research has shown that nurturing self-compassion can improve our psychological health (Gilbert 2010). This has been seen in both healthy and mentally ill individuals (Gilbert and Procter 2006; Neff and Germer 2013). These findings have seen the rise of Compassion Focused Therapy. Compassion Focused Therapy aims to control self-criticism and the harshness with which we often address ourselves. By replacing self-criticism with self-compassion we are able to generate positive emotions (such as warmth and tenderness) within ourselves, as well as promoting non-judgemental acceptance that what we are experiencing is part of the wider human condition.

Our self-compassion derives from our experience of compassion to and from others (Gilbert 2010). Therefore, one therapeutic technique used to nurture self-compassion is developing an image of a compassionate other and then seeing yourself as this person. This involves switching from your perspective to that of the compassionate other. Research suggests that understanding another person's visual perspective is associated with understanding their mental perspective; their intentions, actions and state of mind (Thakkar & Park, 2010).

The purpose of the current study was to establish how effective virtual reality (VR) and mental imagery are as tools to cultivate compassionate experiences with the ultimate aim of promoting self-compassion, reducing self-criticism and reducing shame. We were also interested in the relationship between visual perspective-taking and changes in self-compassion.

During the experiment you "became" the child. From this position you saw yourself, as an adult, responding with compassion. This part of the experiment represents self-compassion. As a child you should still identify with the adult and acknowledge the response given is your own. Thus, you are receiving compassion from yourself. The questionnaires completed before the experimental session will enable us to determine how self-compassionate you are, on average. The questionnaires and tasks you completed just before and just after the compassion scenario will be used to gauge any changes in your self-compassion, self-criticalness and shame as a result of the scenario and their relationship to visual perspective-taking.

It is our hope that your data can further our understanding of self-compassion and how it can be applied to improve our psychological health. Your participation is a valued contribution to this new and influential field of Psychology. We have provided some links below that will take you to two websites pioneering in compassion based research. On these websites you can find more information about self-compassion and also Compassion Focused Therapy, including audio help guides in administering self-compassion:

The Compassionate Mind:	Self-Compassion:
http://www.compassionatemind.co.uk/	http://www.self-compassion.org/

Contact Details

Anneka Holden and Nicola Alden (Trainee Clinical Psychologists)
Email: anneka.holden@ucl.ac.uk or n.alden.12@ucl.ac.uk
or
Prof. Chris R. Brewin (Clinical Psychologist)
Email: c.brewin@ucl.ac.uk

Clinical, Educational & Health Psychology
University College London
Gower Street, WC1E 6BT
London, U.K.

If you have any concerns arising from this experiment please use the contacts above. For additional support and advice about mental health please contact one of the following:

Samaritans
Confidential support for people experiencing feelings of distress or despair.
Phone: 08457 90 90 90 (24-hour helpline)
Website: www.samaritans.org.uk

Rethink Mental Illness
Support and advice for people living with mental illness.
Phone: 0300 5000 927
Website: www.rethink.org

References

Gilbert, P. (2010). Compassion Focused Therapy: A Special Section. *International Journal of Cognitive Therapy 3*(2), 95-96.
Gilbert, P., & Procter, S. (2006). Compassionate mind training for people with high shame and self-criticism: Overview and pilot study of a group therapy approach. *Clinical Psychology and Psychotherapy*, *13*(6), 353-379.
Neff, K. D. & Germer, C. K. (2013). A Pilot Study and Randomized Controlled Trial of the Mindful Self-Compassion Program. *Journal of Clinical Psychology 69*(1), 28-44.
Neff, K. D., & McGehee, P. (2010). Self-compassion and psychological resilience among adolescents and young adults. *Self and Identity*, *9*(3), 225-240.
Thakkar, K. N., & Park, S. (2010). Empathy, schizotypy, and visuospatial transformations. *Cognitive Neuropsychiatry*, *15*(5), 477-500.

Appendix D: Script for the Mental Imagery Intervention

Imagine that an 8 year old girl is sat on a chair opposite you. She has blonde hair in a ponytail and is wearing a pink t-shirt and blue jeans.
She is upset and crying. Her body is hunched up and facing down, she is looking down and holding her hands up to cover her face. Her head moves up and down slightly as she cries.

When you are ready, in your head say the first step of your compassionate response, the validation response. *(Pause)*
Imagine the girl is absorbing what you have said, thinking about it and processing it.
Imagine that she responds by moving her hands down, away from her face but she continues to cry.
When you are ready press the button to continue *(Stop- participant clicks the button to continue)*

In your head say the second step of your compassionate response, the redirection of attention response. *(Pause)*
Imagine the girl is absorbing what you have said, thinking about it and processing it.
Imagine that she sits upright a little, raises her face upwards a little and stops crying.
When you are ready press the button to continue *(Stop- participant clicks the button to continue)*

In your head say the third step of your compassionate response, the memory activation response. *(Pause)*
Imagine the girl is absorbing what you have said, thinking about it and processing it.
Imagine that she responds by sitting upright and lifting her head up to look at you.

When you are ready press the button to continue *(Stop- participant clicks the button to continue)*

Now you are going to imagine the same scenario but from the perspective of the child that was upset and crying. You are looking at your adult-self sitting opposite where you are. Your adult-self is going to say the 3 step compassionate response to you. After you have imagined them saying each step take some time to absorb and respond to what they have said in your imagination. Then, when you are ready press the button to continue.

So imagine that you are the child and you are upset and crying, you can see your adult-self seated opposite you *(pause).* When you are ready imagine your adult-self is saying the first step of the compassionate response, the validation response to you. Hear them saying it to you.

Now imagine the adult is saying the second step of your compassionate response, the redirection of attention response. Hear them saying it to you.

Now imagine the adult is saying the third step of your compassionate response, the redirection of attention response. Hear them saying it to you.

This is the end of the task. When you are ready open your eyes and take off the headphones.

Appendix E: Ethics and Consent Forms

E1. Ethics amendment form and approval

E2. Ethics approval

E3. Consent forms

E1. Ethics amendment form

1	**Project ID Number:** DSD.2013.010	**Name and e-mail address of Principal Investigator:** **Prof. Chris R. Brewin** **c.brewin@ucl.ac.uk**
2	**Project Title: Cultivating Compassion**	
3	**Type of Amendment/s (tick as appropriate)**	
	Research procedure/protocol (including research instruments) ☑ Participant group Sponsorship/collaborators Extension to approval needed (extensions are given for one year) Information Sheet/s ☑ Consent form/s ☑ Other recruitment documents Other Please specify:	
4	**Justification** (give the reasons why the amendment/s are needed): This additional study aims to extend the current research programme investigating nurturing compassion through virtual reality. The effectiveness of Immersive Virtual Reality (VR) in cultivating compassion will be explored in comparison to non-immersive computer VR and mental imagery. The impact of state shame and the relationship between visual perspective taking ability and effectiveness of the intervention will also be assessed. The study will be conducted by two DClinPsy trainees for their major doctoral project.	
5	**Details of Amendments** (provide full details of each amendment requested, state where the changes have been made and attach all amended and new documentation) Comparing delivering Compassion through Virtual Reality with Non-immersive Virtual Reality (video technology) and Mental Imagery (Information Sheet Included) This study will investigate the effectiveness of immersive VR (see point 1 in the original ethics application) in cultivating compassion in comparison to non-immersive computer VR (see point 2 in the original ethics application) and mental imagery. The mental imagery condition is a direct analogue of the immersive VR scenario except that the participant is guided by an audio recording to imagine the scenario described above, including the perspective change, where the participant is encouraged to imagine themselves in the position of the child receiving compassion from the self. The effect of the conditions on trait levels of self-compassion, shame, self-criticism and mood will be assessed. Additionally the relationship between both egocentric and allocentric visual perspective judgements, avatar embodiment and the effectiveness of the conditions (as measured by pre-post intervention change in levels trait self-compassion and self-criticism) will be investigated. All conditions will take place in the UCL VR lab. Healthy participants will be selected who score above average on a measure of trait self-criticism. Those who are eligible to take part will then complete a measure of trait self-compassion. These measures will be administered online. Participants will then be randomly allocated, allowing for gender-balance, to one of the three intervention conditions. Prior to the intervention participants will complete online measures of self-compassion, self-criticism, shame and mood. They will also complete a 2D screening task for egocentric visual rotation and a 3D task to measure allocentric perspective perception and memory. Post intervention measures of self-compassion, self-criticism, shame, mood and experience of the intervention will be administered online. Following the intervention, text message reminders will be sent to participants every other day to encourage them to practice imagining the scenario in the intervention. Two weeks later participants will be asked via email to complete online the same state measures of self-criticism, compassion and	

shame again, as well as measures of vividness, ease of recall and frequency of practice of the imagined scenario. A description of these measures can be seen below.

Attached are Information Sheets and Consent Forms.

We would like to add the following measures:
The Test of Self-Conscious Affect-3; TOSCA-3 (Tangney, Dearing, Wagner, & Gramzow, 2000): This questionnaire measures shame, guilt, pride, and embarrassment. These scales are dispositional measures, and are very frequently used in the social-personality literature to assess shame- and guilt-proneness. The TOSCA-3 consists of 16 scenarios followed by four questions regarding the scenarios (each question corresponding to one of the four subscales). Responses are rated on a 5-point scale.

The Experience of Shame Scale; ESS (Andrews, Qian, Valentine, & Source, 2002): this is a 25 item scale designed to assess four areas of characterological shame: shame of personal habits, manner with others, sort of person (you are), and personal ability; three areas of behavioural shame: shame about doing something wrong, saying something stupid, and failure in competitive situations; and bodily shame: feeling ashamed of (your) body or any part of it. For each of these areas, a question is asked relating to experiential, cognitive and behavioural components of shame. Participants rate each item according to how they have felt in the past year on a 4-point scale where 1 is 'not at all' and 4 is 'very much'.

The State Shame and Guilt Scale; SSGS (Marschall, Saftner, & Tangney, 1994): This is a self-reporting scale of state feelings of shame, guilt, and pride. Fifteen items (five questions correspond to each of the three subscales) are rated on a 5-point scale where 1 is "not feeling this way at all", 3 is "feeling this way somewhat" and 5 is "feeling this way very strongly".

Imagery Vividness. Participants will be asked to report on the extent to which they can (1) hear the voice of the image, (2) see the facial expressions of the image, (3) visualise the gestures of the image, (4) picture the image interacting with them (5) giving compassion and (6) receiving compassion, on a 5-point scale where 1 is "perfectly clear and as vivid as in-person," 2 is "clear and reasonably vivid," 3 is "moderately clear and vivid," 4 is "vague and dim," and 5 is "no image at all, you only 'know.'"

Ease of recall. Participants will be asked 'How easy was it for you to recall the scenario?' This will be measured on a 5-point Likert scale ranging from 0 (very difficult) to 5 (very easy).

Frequency of recall. Participants will be asked 'Over the past two weeks, how often have you recalled the image generated by the scenario?' Participants' responses will be measured on a 7-point scale, where 1 is "never", 2 is "less than once a week", 3 is "once a week", 4 is "twice a week", 5 is "every other day", 6 is "once a day" and 7 is "More than once a day".

Little Man Task (Ratcliff, 1978). This task is designed to measure egocentric visual perspective-taking. Thirty-two stimuli cards are presented by the researcher in a random order. Each card shows a manikin presented in one of four orientations with a black disc marking either the right or left hand of the manikin. There are an equal number of stimuli for each possible presentation. On each trial the participant is required to state which of the manikin's hands is marked with a black disc. Performance will be determined by the total number of correct responses.

Topographical perception task (Hartley et al., 2007). This is a 15 item, concurrent match to sample task which measures perceptual allocentric visual perspective taking. The participant is presented with a ''sample'' image, and simultaneously a four-alternative choice of scenes arranged randomly in a 2x3x2 grid. The participant is given a maximum of 30 seconds to identify the target image that matches the topography of the sample image. Each of the landscapes depicted in the three foil images have been constructed so as to resemble the target in different ways (spatial, configural or elemental differences). No image is repeated. Performance is determined by the total number of correct responses.
Topographical memory task (Hartley et al., 2007). This 15 item, delayed match to sample task measures memory for allocentric visual perspective taking. It is essentially the same as the perception task, except that

	the sample image is presented in isolation for approximately 8 seconds before the four-alternative choice of scenes are shown. Performance is determined by the total number of correct responses.
6	**Ethical Considerations** (insert details of any ethical issues raised by the proposed amendment/s; in the case of adding a new researcher, please confirm in writing that you have discussed ethical issues of the project with this researcher and that you have taken them through the risk assessment form for the project, which they have signed) I can confirm that I have discussed the ethical issues of the project with the Principal Investigator and he has gone through and signed the risk assessment form. 1. Measures: The additional measures are regarded as innocuous for healthy participants. However, participants will be made aware that they do not have to answer any questions that they do not wish to. Additionally, if participants have any concerns after answering the questions they can speak with Prof. Brewin, a clinical psychologist, if they feel the need to.
7	**Other Information** (provide any other information which you believe should be taken into account during ethical review of the proposed changes)
	Declaration (to be signed by the Principal Researcher) • I confirm that the information in this form is accurate to the best of my knowledge and I take full responsibility for it. • I consider that it would be reasonable for the proposed amendments to be implemented. Signature: [redacted] Date: 12/13

E2. Ethics approval

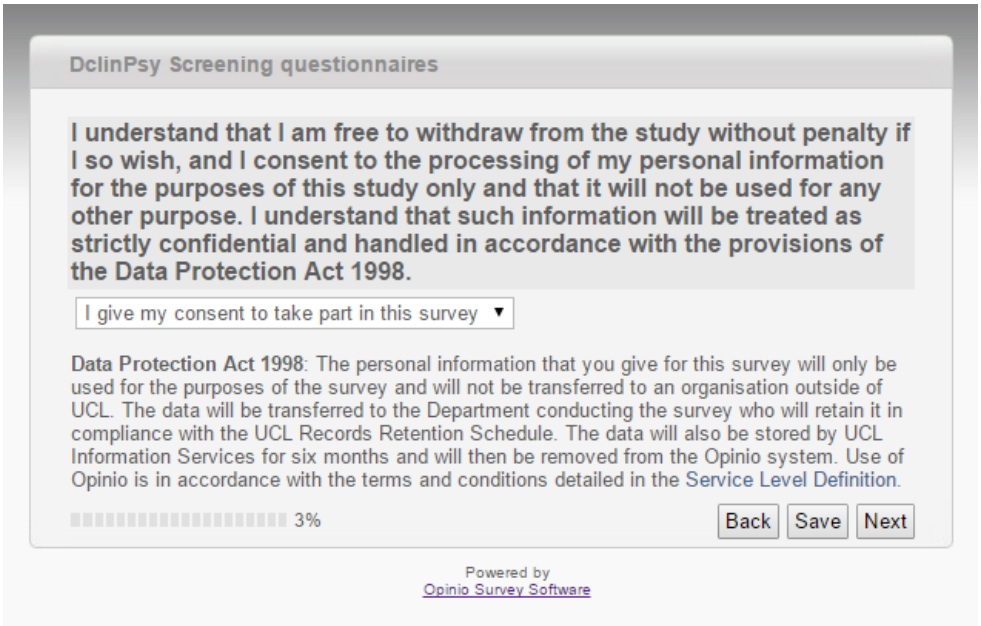
Re: Ethics Docs

Wed 05/02/2014 20:11

Dear Nicola and Anneka,
Many thanks for making these changes. I am happy to approve this. I will send your documents to ethics for archiving, but other than that nothing more needs to be done. Best of luck with this very interesting project!
Best wishes,
Lorna

Dr Lorna Halliday
Lecturer in Developmental Disorders of Communication
Division of Psychology and Language Sciences
University College London
Room 319
Chandler House
2 Wakefield Street
London WC1N 1PF
tel: 44 (0)20 7679 4265 (ext 24265)
email: l.halliday@ucl.ac.uk

E3. Consent forms



Consent Form

Title of project: Approaches to Nurturing Compassion

This study has been approved by the UCL Research Ethics Committee. [Project ID Number: DSD.2013.010]

Investigators: Nicola Alden, Anneka Holden, Prof Chris Brewin, Dr John King

---------- Please read the following carefully ----------

- Thank you for considering taking part in this research. You should only agree to take part after the project has been fully explained to you.

- If you have any questions arising from the information sheet or explanation already given to you, please ask the researchers before you decide whether to join in.

- If you decide at any time during the research that you no longer wish to participate in this project, you can notify the researchers involved and be withdrawn from it immediately and ask to have data about you deleted.

- By signing this document you give your consent to the processing of your personal information, including the data, for the purposes of this research study. You understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.

Participant's Statement

I ……..
agree that the research project named above has been explained to me to my satisfaction and I agree to take part in the study. I have read both the notes written above and the Information Sheet about the project, and understand what the research study involves.

Signed.....…….. Date.....……..

Researcher's Statement

I……
…confirm that I have carefully explained the purpose of the study to the participant and outlined any reasonably foreseeable risks or benefits.

Signed.....…….. Date.....……..

Appendix F: Data for the VREQ and MIEQ

Table F1

Medians, Interquartile Ranges and Difference Between Corresponding Statements for the VREQ (IVR Condition) and MIEQ (MI Condition)

VREQ statement (*n* = 20)	**MIEQ statement (*n* = 20)**	**IVR**		**MI**		**U**	***p*** [a]	**Effect size *r***
		Median	**IQR**	**Median**	**IQR**			
Section A: Presence								
During the experience my strongest feeling was that I was in the lab taking part in an experiment as opposed to a room where there was a child crying	During the experience my strongest feeling was that I was in the lab taking part in an experiment as opposed to a room where there was a child crying	1.00	-0.75, 2.00	2.00	1.00, 2.00	157.50	.22	
The appearance of the lab (walls, door, curtain, carpet) I saw while wearing the helmet kept reminding me I was in a simulation as opposed to helped me to forget I was in a simulation	The headphones kept reminding me I was in a simulation as opposed to helped me to forget I was in a simulation	0.50	-1.00, 1.75	1.50	0.25, 2.00	135.50	.08	
I was aware of the mirror and I almost never looked at it as opposed to I looked at it all the time		2.00	1.00, 2.00					
Section B: Body Ownership and Agency as Adult								
I felt as if the body I saw when I looked down was my own body.		1.00	0.25, 2.00					
I felt as if the body I saw when I looked in the mirror was my own body.		1.00	-1.00, 2.00					

Table continues

VREQ statement (*n* = 20)	MIEQ statement (*n* = 20)	IVR		MI		U	*p* [a]	Effect size *r*
		Median	IQR	Median	IQR			
I had the feeling that I was looking at myself in the mirror rather than looking at someone else.		1.00	-2.00, 2.00					
The movement of the adult's body responded to the movements of my real body.		2.00	1.25, 3.00					
My talking to the child had a positive effect on her state.	My talking to the child had a positive effect on her state.	2.00	1.00, 2.00	2.00	1.00, 2.00	178.00	.51	
While I was talking to the child I was concerned about forgetting my lines.	While I was talking to the child I was concerned about forgetting my lines.	1.00	-0.75, 3.00	0.00	-2.00, 1.00	122.00	.03*	.34
When I spoke to the child I really meant what I said, as if it had come from my heart.	When I spoke to the child I really meant what I said, as if it had come from my heart.	1.00	-1.00, 2.00	2.00	0.25, 2.00	164.00	.32	
When I spoke to the child I had the impression I was just delivering someone else's lines	When I spoke to the child I had the impression I was just delivering someone else's lines	1.00	-0.75, 1.00	0.50	-1.50, 1.00	192.50	.83	
How much did you feel like you had two bodies? (Reverse scored)		2.00	-1.00, 2.00					
How much did you feel like the child was aware of your presence?	How much did you feel like the child was aware of your presence?	1.50	1.00, 2.00	2.00	1.00, 2.75	165.00	.32	
How much did you feel like you were in control of the adult avatar?		2.00	1.25, 3.00					
Section C: Body Ownership and Agency as Child								
I felt as if the body I saw when I looked down was my own body.		1.00	-0.75, 2.00					
I felt as if the body I saw when I looked in the mirror was my body.		1.00	-0.75, 2.00					

Table continues

VREQ statement (*n* = 20)	MIEQ statement (*n* = 20)	IVR		MI		U	*p* [a]	Effect size *r*
		Median	IQR	Median	IQR			
The movement of the child's body responded to the movements of my real body.		2.00	2.00, 3.00					
I felt myself to be in the child's role.	I felt myself to be in the child's role.	1.00	0.25, 2.00	1.50	0.25, 2.00	191.50	.81	
How much did you feel like you were in control of the child avatar?		2.00	2.00, 3.00					
When I saw the adult giving compassion, it was easy to recognise myself as that adult.	When I saw the adult giving compassion, it was easy to recognise myself as that adult.	0.00	-2.00, 2.00	2.00	-0.50, 2.00	138.50	.09	
When I saw the adult giving compassion, it was easy to recognise myself in her body movements.		1.00	-1.00, 2.00					
When I saw the adult giving compassion, it was easy to recognise myself in her voice.	When I saw the adult giving compassion, it was easy to recognise myself in her voice.	2.00	1.00, 3.00	1.00	-1.00, 2.00	122.00	.03*	.34
I had the feeling I was giving compassion to myself.	I had the feeling I was giving compassion to myself.	1.00	-1.75, 2.00	1.00	-1.00, 2.00	168.50	.38	
How much did you feel like you had two bodies? (Reverse scored)		0.00	-1.75, 2.00					
I felt comforted by the adult avatar.	I felt comforted by the adult avatar.	-0.50	-2.00, 1.00	2.00	0.00, 2.00	80.50	.001**	-.52
I felt reassured by the adult avatar.	I felt reassured by the adult avatar.	-1.00	-1.75, 1.00	1.50	0.25, 2.00	90.00	.002**	-.48
I felt threatened by the adult avatar.	I felt threatened by the adult avatar.	-2.00	-3.00, 0.75	-2.50	-3.00, -2.00	163.50	.29	
I felt that I was critical of the adult avatar.	I felt that I was critical of the adult avatar.	1.00	0.00, 2.00	-2.00	-3.00, 1.00	117.50	.02*	.36

Note. Each statement was rated on a 7 point Likert scale ranging from -3 to 3 where -3 = 'not at all' and 3 = 'very much so', except for: the first statement where -3 = '…the lab taking part in an experiment' and 3 = '…a room where there was a child crying'; the second statement where -3 = 'kept reminding me I

was in a simulation ‘and 3 = ‘helped me to forget I was in a simulation’; the third statement where -3 = ‘I almost never looked at it’ and 3 = ‘I looked at it all the time’.

VREQ = Virtual Reality Experience Questionnaire; MIEQ = Mental Imagery Experience Questionnaire; IVR = immersive virtual reality; MI = mental imagery

[a] Significance level: $p < .05$ or Bonferroni correction $p < .003$

***** Significant at $p < .05$

******Significant at $p < .003$