BABEŞ-BOLYAI UNIVERSITY
FACULTY OF PSYCHOLOGY AND EDUCATIONAL SCIENCES
DOCTORAL SCHOOL "EVIDENCE-BASED PSYCHOLOGICAL ASSESSMENT AND INTERVENTIONS"

Ph.D. THESIS SUMMARY

A MULTINIVELAR ANALYSIS OF ATTENTIONAL BIAS EVALUATION AND VIRTUAL REALITY MEDIATED ATTENTIONAL BIAS MODIFICATION INTERVENTIONS IN ANXIETY AND DEPRESSION

AUTHOR: Ph.D. CANDIDATE FODOR LIVIU - ANDREI
SCIENTIFIC ADVISOR: PROFESSOR Ph.D. SZAMOSKÖZI ȘTEFAN

CLUJ-NAPOCA
2022

ACKNOWLEDGEMENTS

This thesis is the result of the interaction between my own effort and the participation of several valuable people. As a result, in the following lines, I'd want to thank those who helped me grow personally and professionally, and without whom I couldn't have completed this work. Therefore, in the beginning, I would like to express my gratitude to Professor Szamosközi Ștefan for supervising my thesis. His academic supervision helped me to gain a better insight into various important methodological and life-related aspects. Next, I would like to thank Professor Daniel David for facilitating and encouraging my research in the virtual reality field, and without whom this work would have been significantly hampered. My gratitude also extends to Ioana Cristea, for her mentorship and patient guidance through the field of meta-research. Her scientific outlook had a decidedly positive impact on my manner of approaching and understanding scientific research. I am also grateful to Professor Aurora Szentagotai, Professor Anca Dobrean, Assisstant Professor Silviu Matu and all the members on the Department of Clinical Psychology and Psychotherapy at Babeș-Bolyai University, for their insightful remarks and support, which aided in the development of the quality of this thesis. Special thanks go to my office colleagues Raluca Georgescu and Ruben Nechifor for their unconditional support over the years. Last but not least, I would like give thanks to my family, friends and to Laura, on whose support and reassurance I could count at any time during the completion of the Ph.D. program.

Notes._____

(1) This is to certify by Fodor Liviu-Andrei that:

(a) The thesis includes the original research work of Fodor Liviu-Andrei towards the Ph.D.;

(b) Parts of the thesis have been already published, in press, or submitted for publication; appropriate citations for these publications were included in the thesis. Other co-authors have been included in the publications, if they contributed to the exposition of the published text, data interpretation etc. (their contribution was clearly explained in the footnotes of the thesis);

(c) The thesis was written according to the academic writing standards (e.g., appropriate scientific acknowledgements and citations have been made in the text by the author of the thesis). All the text of the thesis and its summary was written by Fodor Liviu-Andrei who assumes the all responsibility for the academic writing; also:

- A software was used to check for the academic writing (see at http://www.plagiarism-detector.com/); the thesis has passed the critical test;
- A copy of the research dataset/database was delivered at the Department/Graduate School.

Signature for certifying the Notes: Ph.D. candidate Fodor Liviu-Andrei

(2) All the Tables and Figures are numbered within the corresponding chapter or subchapter of
the thesis

TABLE OF CONTENTS

CHAPTER 1. THEORETICAL BACKGROUND

1.1. Introduction and Research Problem

1.1.1. General Framework

The overall objective of the present thesis is represented by our endeavour to evaluate the effectiveness of novel, technologically-augmented interventions in evaluating and modifying attentional biases, with the scope to reduce anxious symptomatology. More specifically, through the present work we intended to take a step forward towards a better understanding of evidence-based virtual reality augmented attention bias modification interventions and how these interventions can be employed to alleviate anxious symptomatology.

In this sense, the present thesis was built around four chapters. In the first chapter, we described the theoretical basis and the research limitations that exist in the field. The second chapter describes the specific objectives of the doctoral thesis and the research methodology that was employed in order to meet those objectives. The third chapter details the original research that was conducted, as well the results that were obtained. The fourth chapter we presented the conclusions that were derived from the original research, as well as the theoretical and practical implications. Lastly, we detailed the inherent limitations of the original research and we suggested possible avenues for future research.

1.1.2. Understanding executive functioning and cognitive biases in general, with a focus on attentional bias

Anxiety and depressive disorders represent the most prevalent mental health disorders, with 3.94% of the world population being affected by an anxiety disorder and 3.59% of the world population being affected by a depressive disorder (Dattani et al., 2021). More specifically, between 10% - 20% of adults will seek professional mental health services in any given 12 month period, in relation to a anxious or depressive disorder episode, with more than 50% of them being affected by comorbid second anxiety or depressive disorder episode (Hirschfeld, 2001). The effects of the high rate of comorbidity between anxiety and depression cannot be understated. In general, people that have comorbid anxiety and depressive disorders tend to have a greater illness severity, a higher chronicity, and significantly greater impairment in work functioning, psychosocial functioning, and quality of life (Olfson et al., 1997, Brown et al., 1996, Kessler et al., 1998, Sherbourne et al., 1996, apud Hirschfeld, 2001).

To survive, all individuals must be ready to adapt to ever-changing environments. The term “executive function” refers to the ability to adapt by regulating reflexive reactions to current stimuli, in order to achieve goals that require the manifestation of complex behaviours. Executive function is thought to be a set of higher order cognitive abilities that allow people to plan for the future, exercise self-control, and successfully complete goal-directed behaviour. According to research conducted in this area, some of the cognitive processes that are included under the umbrella term of executive function are attention, impulse/inhibitory control, working memory, decision making, cognitive flexibility, planning (Baddeley, 1998, Robbins, 1996, Stuss & Alexander, 2000). Directly related to attention as an executive function, attentional control is defined as the ability of an individual to pick what they pay attention to and what they ignore and is mediated primarily by the anterior cingulate cortex. One of the most fundamental assumptions of attentional control theory is that knowing how anxiety affects attentional performance requires understanding the effects of anxiety on attentional processes. One suggestion is that worry can impair

attentional control via a process that alters the balance between the goal-directed and stimulus-directed attentional systems, with the stimulus-driven attentional system having a far greater effect (Eysenck et al., 2007). More precisely, deficits in neurocognitive functioning are associated with anxious/depressive symptomatology, with an abundance of evidence highlighting the role of impairments in attentional control (Eysenck et al., 2007; Pacheco-Unguetti et al., 2011; Rock et al., 2014). The evidence that supports this hypothesis comes from research focused on cognitive biases, more specifically from studies investigating attentional biases (Eysenck et al., 2007). Specifically, it has been shown that there is a strong association between attentional control and attentional biases in anxious adults (Derryberry & Reed, 2002). Moreover, attentional control acts as a moderator for the relationship between attentional biases and anxious symptomatology (Campbell & Kertz, 2019; Susa et al., 2012).

The concept of cognitive bias was firstly introduced in the early 70's, circumscribing systematic errors of thought that people are prone to commit when having to process and interpret the informational stimuli surrounding them (Tversky & Kahneman, 1974). While the human brain can process a fair amount of information, it must inevitably rely on shortcuts or simplifications in order to accommodate the volume of information it is confronted with on a daily basis. Tversky & Kahneman (1974) described three such shortcuts in their article, which they called "heuristics" or "cognitive biases": representativeness, availability of instances or scenarios, and adjustment from an anchor. Representativeness is usually employed when people have to judge if an event has a certain probability of happening, while comparing that event with a known mental prototype. The availability of instances or scenarios heuristic is a mental shortcut, meaning that people are relying on fast / immediate mental examples when having to assess a certain informational input. The adjustment from an anchor heuristic describes the phenomenon in which people are influenced by a reference point / anchor when having to make certain decisions. While such cognitive biases can lead to rapid, economical and effective decision making in people, more often than not they are responsible for distorted perceptions, illogical interpretations and inaccurate decision making (Kahneman & Tversky, 1972). Since then, more than 180 cognitive biases have been identified.

The most widely studied cognitive biases, which have been proposed to have causal and disorder-maintaining effects in anxious and depressive symptomatology are attention and interpretation biases.

Attention bias is defined as "the tendency to prioritize the processing of certain types of stimuli over others" (Azriel & Bar-Haim, 2020). More specifically, people can be perceptually confronted with potentially unlimited stimuli from the environment and, because the human cognitive system has limited resources, attention is directed towards certain types of stimuli, while discarding others. When this process takes place in relation to threatening stimuli, it is known as threat-related attention bias. The exact definition of threat-related attention bias has been proposed as "the tendency to prioritize the processing of potential threats over benign stimuli" (Azriel & Bar-Haim, 2020). In other words, when individuals are confronted with a threatening stimulus, they tend to prioritize it, even if other types of stimuli are present in the environment and competing for attention, such as neutral or positive stimuli.

The scientific evidence for the link between anxious symptomatology and attention bias to threat has largely been derived from research approaches involving the dot probe task (MacLeod et al., 1986). This computerised task consists of a large

number of trial repetitions. In each trial, the participant observes a series of events, namely: 1) a fixation cross that has the role of directing the participant's gaze towards the computer screen; 2) after the fixation cross disappears from the screen, for a short duration (usually 500 milliseconds), a pair of stimuli are presented on the computer screen, one threatening and one neutral (or positive in some research areas); 3) after the stimuli disappear from the screen, a target probe is presented in the location in which one of the stimuli was presented. The participant has to react as fast as he/she can to the target probe, usually by pressing a key and the allocation of attention is determined as a function of the time needed to react to the target probe. The majority of the studies found that participants with anxious symptomatology respond faster when the probe replaces the threatening stimulus, than when the probe replaces the neutral stimulus (Mogg & Bradley, 1998).

While such a process (an attentional bias to threat) can have certain evolutionary advantages, it can also lead to an oversensitivisation to detecting threat. There is ample evidence for this phenomenon happening in individuals with anxious symptomatology (Williams et al., 1997). Attentional bias has been considered to have an active role in initiating and maintaining anxious symptomatology (Koster et al., 2004), while also being responsible for other changes such as increasing the frequency, intensity and duration of anxious symptomatology (Azriel & Bar-Haim, 2020). Moreover, it has been demonstrated that having an attentional bias to threat leads to lesser improvements following cognitive-behavioural therapy (Campbell & Kertz, 2019).

Given the causal role of the attentional bias in anxiety disorders, a procedure called attention bias modification training has been developed in order to reduce attentional bias towards threat-related stimuli and, through this mechanism, to reduce anxious symptomatology. More precisely, by taking an alternative route to classical cognitive-behavioural therapies which aim at changing the automatic thoughts through cognitive restructuring and thus reducing anxious symptomatology, the attention bias modification training procedures aim at reducing anxious symptomatology by changing the attentional bias towards threat instead of changing the automatic thoughts (David et al., 2013). Thus, the underlying schema does not produce negative automatic thoughts and the anxious symptomatology is reduced (David et al., 2013). The attention bias modification training is a computerized intervention and makes use of the dot probe paradigm described above for bias modification, while aiming at automatically training alternative way of processing information. The dot probe is modified in such a way that when the researchers wish to simply evaluate attentional bias, the probe replaces the negative and neutral (or positive) stimuli with equal frequency, this modality assuring that no bias modification takes place. A bias index score can be calculated by subtracting the mean reaction time of the participants toward disgust faces from the mean reaction time towards neutral faces. When bias modification training away from threatening stimuli and towards neutral (or positive) stimuli is intended, then the dot probe is modified in such a way that the probe replaces the neutral stimuli in 80% - 100% (in some cases 20% of trials employ neutral-neutral stimulus pairs in order to mask trial contingency) of the total number of trials. Previous research has employed between 40 and 750 trials per session, with a number of 128 trials being the most common (Bar-Haim, 2010). The total number of sessions were between 1 and 15, most studies employing either one or eight sessions (Bar-Haim, 2010).

The advantages of such an attention bias modification intervention are immediately obvious. Firstly, the intervention is very easy to deliver, the only equipment needed being a desktop computer or a laptop. The attention bias modification task is relatively easy to build and configure locally using a software such as PsychoPy, Inquisit or Pebl or it can be implemented online via the PsychoPy software or via classical web programming. Moreover, the procedure is wholly automated, with no need for the therapist to intervene during the procedure. These advantages also make attention bias modification procedures ideally suited for being used as an addition to classical therapeutic interventions or as therapeutic homework.

There are also a number of disadvantages that current attentional bias evaluation methods and attention bias modification procedures are plagued with, among which the most relevant are the mixed results when it comes to therapeutic effectiveness for attention bias modification, and the fact that the procedures themselves are monotone, non-interactive.

Firstly, there are mixed findings in the literature when discussing the effectiveness of attention bias modification procedures. With regard to attention bias modification interventions for anxiety disorders, a number of meta-analyses reported small, frequently non-significant, symptom reductions compared to control conditions (Cristea et al., 2015; Heeren et al., 2015), while others reported significant effects of larger magnitude (Linetzky et al., 2015; Price et al., 2016). Moreover, the interventions have only been compared in a pairwise manner, there being no investigation in which cognitive bias interventions for anxious symptomatology are compared simultaneously.

Secondly, it has been suggested in previous research that the attentional bias evaluation and attention bias modification procedures are monotone and can pe perceived as boring by the research participants, as the number of trials increases (Van Ryckeghem et al., 2018). To counter this, a number of solutions have been proposed. On one hand, one suggested solution would be to augment the interest towards the task by employing motivational elements such as a reward for appropriate performance or introducing gamification elements such as a points system (Dennis & O'Toole, 2014; Karoly & Crombez, 2018). On the other hand, an alternative solution would be to perform the attention bias modification training by using a similar to real-life context. For example, in the case of attention bias modification for pain, it has been proposed to use cues of actual pain stimuli, rather than semantic representation of pain (i.e., words that are often associated with the sensation of pain or images suggestive of pain) presented in a safe context (Karoly & Crombez, 2018; Van Ryckeghem & Crombez, 2014). In the case of anxiety disorders, an adaptation of the aforementioned proposal would not be very feasible. An alternative approach would be to change the medium primarily, rather than the stimuli, namely to implement the training in such a way that it can be performed in a augmented / virtual reality medium that can be built and customised to resemble real-life contexts and scenarios.

1.1.3. Understanding virtual reality interventions

In the 1950s and 1960s, numerous major inventions paved the way for what is today known as virtual reality. Morton Heilig created the Sensorama in 1957, with the goal of engaging all of the user's senses with particular components such as fragrance generators and vibrating chairs to provide a complete multi-sensory experience (Dinh et al., 1999). For military training, the Philco Corporation developed Headsight in 1961, the first head-mounted display with motion tracking and dual monitor displays (Comeau, 1961). Ivan Sutherland created the Ultimate display in

1965, which had the first computer-generated interface, allowing people to interact with VR in real time (Sutherland, 1965). When Jaron Lanier invented the word "virtual reality" in 1989, the concept of VR became more defined, and VR became more prevalent in research and mental treatment (Conn et al., 1989; Gorini & Riva, 2008).

Since then, virtual reality interventions have been used in for a broad range of therapeutical targets, either as standalone interventions or as components of composite therapeutic plans. Empirical research was conducted to evaluate the efficacy of virtual reality interventions for mental health symptomatology such as fear of flying (Hodges et al., 1996; Wiederhold & Wiederhold, 1998; North & Rives, 2003), fear of driving (Kaussner et al., 2020), fear of heights (Rothbaum, Hodges, Kooper, et al., 1995; Rothbaum, Hodges, Opdyke, et al., 1995; North et al., 1996a), agoraphobia (North, North, & Coble, 1995b; North, North, & Coble, 1996b), claustrophobia (Booth & Rachman, 1992; Botella et al., 1998), fear of public speaking (Harris, Kemmerling, & North, 2002), autism spectrum disorder (Strickland, 1996), body experience in eating disorders (Riva, 1997), posttraumatic stress disorder (Rothbaum et al., 1999; Rizzo et al., 2010), obsessive-compulsive disorder (North & North, 2000), attention deficit hyperactivity disorder (Rizzo et al., 2000). Specifically, virtual reality therapy for anxiety disorders has primarily arisen as a practical alternative to imaginal and in vivo exposure, treatments that, despite their undeniable effectiveness, are not usually adopted due to being perceived by both patients' (Garcia-Palacios et al., 2007) and even therapists' (Schumacher et al., 2017) as invasive. Moreover, the research that was conducted until now to investigate the efficacy of virtual reality interventions has revealed promising results. One of the first meta-analysis conducted in this sense (Parsons & Rizzo, 2008) has indicated that the effect sizes in favour of virtual reality as compared to agglutinated control groups were very high, from 0.87 for post-traumatic stress disorder to 1.79 for panic disorder with agoraphobia. Another pioneer meta-analysis (Powers & Emmelkamp, 2008) has revealed that 1) therapies involving virtual reality were more efficient than control groups (all types mixed together), with a large effect size of d = 1.11, 2) therapies involving virtual reality were more efficient than control groups with regard to general distress measures (g = 0.5), cognitive measures (g = 1.30), behavioral measures (1.27) and psychophysiology measures (g = 0.68). Finally a more recent and rigorous meta-analysis (Opriș et al., 2012) has identified 1) a large effect size in favour of virtual reality interventions as compared to wait-list (d = 1.12), a result that had stability in time (at follow-up) and 2) the fact that there is a direct connection between dosage and response, namely that the number of virtual reality sessions is a statistically significant moderator of the effect size, namely the effect size increases linearly with the increase of the number of sessions. In conclusion, it would seem that virtual reality interventions are effective compared to "passive" control conditions and at least as effective as "active" control conditions. Virtual reality interventions come with a series of advantages and disadvantages however, that have to be weighted before being employed in a treatment scheme. Among the advantages, without being exhaustive, we can mention the fact that simpler virtual reality environments can be developed and implemented on smartphones, thus circumventing the need for very expensive computer configurations. Another advantage is represented by the fact that virtual reality therapy has the potential to have high ecological validity, be very immersive and immersion (i.e., quantity and quality of sensory data that is perceived from the virtual reality environment) directly determines user presence (Cummings & Bailenson, 2016). Presence has been defined as the patient's sense of "being there" in the virtual environment (Slater & Wilbur,

1997), a construct that has been shown to be associated to therapeutic effectiveness (Price, 2011). Moreover, virtual reality applications have the potential to give, as an output, objective metrics of treatment effectiveness, such as total scores, points, number of errors and so on, providing alternative avenues of treatment efficacy exploration, beyond the classical clinician-rated and/or self-report instruments. Among the disadvantages of virtual reality therapies, again, without being exhaustive, we can mention cyber sickness (i.e., nausea induced by the virtual environment, caused by motion in the virtual environment) and cost-effectiveness, in the sense that complex virtual reality environments (in terms of visual / auditory attributes and interactivity) can be run only on relatively expensive hardware configurations.

1.1.4. Virtual reality attention bias modification

When speaking about attention bias modification procedures in virtual reality, the literature is extremely sparse. To date, there are only two studies that embedded attention bias modification procedures in virtual environments (Urech et al., 2015; Ma et al., 2019). While providing significant and very informative results with regard to virtual reality-based attention bias assessment and modification, these studies have a series of limitations, respectively 1) the first study was conducted as a proof-of-concept study, adopting a pre-post intervention design, without employing a control group (Urech et al., 2015) and 2) the second study employed no "classical" intervention control group, all groups experiencing VR immersion, the aim of the study being to discern the efficacy of different stimuli dimensionalities. This state of affairs illustrates clearly why more research is needed in the field of virtual reality-based attention bias assessment and modification.

1.1.5. Summary and Concluding Remarks

Anxiety and depression represent major global issues, with a profound clinical, economic and social impact, with more than 7% of the world population suffering of one of these mental health problems at any given time (Dattani et al., 2021). We know that deficits in neurocognitive functioning are associated with anxious/depressive symptomatology, especially impairments in attentional control (Eysenck et al., 2007; Pacheco-Unguetti et al., 2011; Rock et al., 2014), as evidenced by studies in the field of cognitive biases, more specifically from studies investigating attentional biases (Eysenck et al., 2007). We also know that the attentional bias has been considered to have an active role in initiating and maintaining anxious symptomatology (Koster et al., 2004), while also being responsible for other changes such as increasing the frequency, intensity and duration of anxious symptomatology (Azriel & Bar-Haim, 2020). Moreover, virtual reality-based assessment and interventions have matured enough, with a significant corpus of research being conducted in this direction. However, there still are some questions that remain and areas of research that merit further investigation.

Firstly, meta-research conducted until now on the efficacy of virtual reality interventions for anxiety has some limitations, one of the most important ones being that comorbidities were not taken into account when investigating treatment efficacy, or it well known that the presence of depressive symptoms is associated with worse treatment outcomes (Kalin, 2020).

Secondly, with regard to cognitive bias modification interventions for anxiety disorders, the findings of meta-research conducted in order to investigate their efficacy are mixed, some research reporting small or non-significant effects (Cristea et al., 2015; Heeren et al., 2015), while others reporting larger effect sizes (Linetzky et al., 2015; Price et al., 2016). As in the case of meta-research of virtual reality-based

interventions, no study took comorbidities into account and the intervention were compared only in a pairwise manner, with some possible combinations or pairings of treatments never being investigated.

Thirdly, even if research regarding the use of attentional bias assessment has matured, it's not very clear if this type of assessment is effective at discriminating between anxious / depressed individuals and healthy controls and thus, being fit to be potentially employed in the diagnosis process.

Fourthly, the research literature on employing virtual reality in modifying attentional biases is extremely sparse, with only two studies breaching this avenue of research, none of which used a control group that would give more insight into the efficacy of virtual reality-based attention bias modification.

CHAPTER II. RESEARCH OBJECTIVES AND OVERALL METHODOLOGY

Through the present thesis, we aimed at addressing a number of methodological objectives related to the virtual reality mediated attentional bias evaluation and attentional bias modification interventions.

The first major goal of the present research was to investigate the efficacy of virtual reality-based interventions, as compared to passive or active control conditions, for anxious symptomatology and comorbid depressive symptomatology. Given that the empirical evidence in this sense has some limitations and there have been no updates in this sense for a significant amount of time, there is a need for updated research in the form of a quantitative synthesis. For this objective, we conducted a updated meta-analysis of 39 studies, comprising 52 direct comparisons between virtual reality-based interventions and passive or active control conditions, while also investigating potential moderators of the effect size, study quality, publication bias and attrition rates (Study 1).

The second major goal of the present research was to investigate the comparative efficacy of cognitive bias modification interventions with regard to anxious symptomatology, depressive symptomatology, comorbid anxious symptomatology and comorbid depressive symptomatology. Given the fact that 1) the interventions have only been compared in a pairwise manner in individual studies, and there being no investigation in which cognitive bias interventions for anxious symptomatology are compared simultaneously in a meta-research framework, and 2) no intervention effects on comorbidities being previously investigated, there is a need for this literature gap to be filled. In this sense, we conducted a network meta-analysis of 85 trials, 65 on anxiety and 20 on depression. (Study 2).

Given the fact that the results that are found in the literature with regard to the discriminatory power of attentional bias evaluation methods are mixed for anxious / depressive individuals and healthy controls, in Study 3 we aimed to pursue this research goal. In order to achieve this, we conducted a study in which anxious / depressive individuals and healthy controls performed an attentional bias evaluation task. We aimed at comparing response times between healthy individuals and individuals with clinic/subclinical levels of anxious/depressive symptomatology, with the goal of establishing if the attentional bias evaluation method based on the dot-probe is efficient in discriminating between affected and healthy individuals (Study 3).

The fourth major goal of the present research, given the scarcity of the literature in this sense, was to investigate the efficacy of a virtual reality-based attention bias modification procedure, as compared with the classical computerized version, in modifying attentional bias and reducing anxious symptomatology. In this sense, we conducted a pilot randomized controlled trial in which we randomized participants either to the virtual reality-based or computerized attention bias modification intervention. We also investigated possible adverse effects, virtual system usability, the level of presence induced by the virtual environment and stress/perceived mental load (Study 4).

Figure 1. The schematic structure of the Ph.D. project

CHAPTER III. ORIGINAL RESEARCH

3.1. Study 1: The effectiveness of virtual reality based interventions for symptoms of anxiety and depression: A meta-analysis[1]

3.1.1. Introduction

Virtual reality (VR) has garnered significant attention as a cost-effective tool for delivering psychological treatments (Freeman et al., 2017). Virtual reality exposure (VRE) in particular is considered an effective treatment for several anxiety disorders (David et al., 2013), on par with in vivo exposure/IVE (Gerardi et al., 2010; Opriş et al., 2012), though doubts were expressed about the quality of this evidence (Meyerbröker & Emmelkamp, 2010).

While many narrative reviews and commentaries focused on VR interventions, only three systematic reviews with meta-analyses examined their efficacy in randomized controlled trials/RCTs (McCann et al., 2014; Opriş et al., 2012; Turner & Casey, 2014) and they present certain shortcomings. Included trials were published through 2014 the latest, and many more trials have been conducted since, given VR technology has become more accessible. Outcomes other than anxiety were scarcely analyzed, though data on some of these has been accruing. The effects of VR interventions on treatment attrition remained unclear, with some speculation of possible superiority (Botella et al., 2015; Freeman et al., 2017; Meyerbröker & Emmelkamp, 2010), but no assessment in a meta-analysis.

Only one meta-analysis (Turner & Casey, 2014) considered heterogeneity between effect sizes (ESs), but did so only descriptively, without providing a quantification. Assessment of quality (McCann et al., 2014; Turner & Casey, 2014) relied on mixed and potentially inadequate tools that included items not linked to any type of trial bias (e.g., treatment fidelity) (Armijo-Olivo et al., 2013), thereby potentially confounding the relationship between study quality and treatment effects. Only one meta-analysis (Turner & Casey, 2014) considered publication bias, with conflicting results between the assessment methods used (Egger's test and fail-safe N). Moreover, many VR trials are conducted on a small number of participants, which exposes meta-analyses to "small study effects" (Sterne et al., 2000), the notion that smaller studies show different, often larger, treatment effects than large ones. Few potential moderators were examined, with generally contradictory results regarding treatment intensity, or the type of comparison group. One yet uninvestigated potential moderator regards the involvement of developers of VR tools and interventions in the trials, as these are often for-profit developments.

Consequently, we report a meta-analysis for the effectiveness of VR-enhanced interventions in RCTs, for symptoms of anxiety and depression, as well as treatment attrition, along with assessment of risk of bias, heterogeneity, and potential moderators.

[1] This study has been published
Fodor, L. A., Coteț, C. D., Cuijpers, P., Szamoskozi, Ștefan, David, D., & Cristea, I. A. (2018). The effectiveness of virtual reality based interventions for symptoms of anxiety and depression: A meta-analysis. Scientific Reports, 8(1), 10323. https://doi.org/10.1038/s41598-018-28113-6

3.1.2. Methods

Identification and selection of studies

A literature search of PubMed, PsycInfo, EMBASE and Cochrane Central Register of Controlled Trials databases was conducted through May, 2015, updated in March, 2016 and subsequently August 2017, using the keywords "virtual reality", "therapy", "exposure", "intervention", "treatment" and a filter for randomized trials (Supplementary Method). We also searched the references from the most recent systematic reviews and meta-analyses.

Studies were included if they were a) RCTs comparing b) a VR-enhanced intervention to a control or an active psychological intervention for c) adults, d) measuring outcomes related to depression and anxiety, and e) published in peer-reviewed journals. We included studies comparing a VR-enhanced condition with controls (e.g., waitlist, placebo, treatment-as-usual) or active conditions not employing VR. Similarly to Turner & Casey (2014), the latter were defined as established interventions involving active, psychologically therapeutic mechanisms of action (e.g., CBT, IVE). No language restrictions were employed. One researcher screened all abstracts and full-texts of RCTs were recovered. Two independent researchers independently examined full-texts and selected eligible RCTs. Disagreements were resolved by discussion and consultation with a third author until consensus was reached.

Risk of bias and data extraction

We used four criteria from the Risk of Bias (RoB) assessment tool, developed by the Cochrane Collaboration (J. P. T. Higgins et al., 2011), which assesses possible sources of bias in RCTs. The following domains were rated: a) the adequate generation of allocation sequence, b) the concealment of allocation to conditions, c) the prevention of knowledge of the allocated intervention (blinding of assessors) and d) the adequate addressment of incomplete outcome data. Blinding of assessors was rated as low risk if the trial described proper methods of ensuring it or if all relevant outcome measures were self-report, thus not requiring the direct interaction with an assessor. This choice was made as we expected most outcomes to be reported on self-report scales, and there is currently no standard as to how to rate these in terms of blinding. Domain d) was assessed as low risk if all randomized participants were included in the analysis, either through the use of an intent-to-treat (ITT) approach or when complete data was available. We also computed an overall RoB score for each study by awarding 1 point for each bias source rated as low risk.

We extracted a series of variables from the included studies, detailed in Table 1 for further use in moderator analyses. Details about the interaction with the virtual environment were extracted from the methods sections describing the intervention or the technology used. For each trial, we noted which elements the interaction with the VR environment relied upon (e.g., visual, sound, haptic) and (2) whether or not the authors had explicitly assessed sense of presence or immersion in the trial with validated or ad hoc instruments. We also quantified the first component by tabulating the number of interaction elements each study employed, as a very crude indicator of the degree of interaction.

The involvement of a developer was coded using the information available in each trial, at the section of the method that described the VR therapy package used. If authors of the VR package were not listed in the original article, we independently searched the web for the specific VR program or package used in order to identify its authors. Risk of bias assessment and data extraction were performed by two

independent researchers and disagreements were discussed and resolved until consensus was reached.

Meta-analyses

We computed and pooled the individual ESs with Comprehensive Meta-Analysis (CMA version 3.3.070) and Stata (Stata SE, version 15).

For anxiety and depression, we calculated the standardized mean difference (SMD) at post-test and follow-up, by subtracting the mean score of the comparison group (control or active treatment) from the mean score of the VR-enhanced group, and dividing the result by the pooled standard deviation of the two groups. Positive SMDs thus reflect superiority of the VR-enhanced condition. We report the indicator corrected for small sample bias (Hedges et al., 1985), Hedges' g. We also transformed the SMD into number needed to treat (NNT), using the formula of Kraemer & Kupfer (Kraemer & Kupfer, 2006). The NNT represents the number of patients that would have to be treated to generate one additional positive outcome (Laupacis et al., 1988).

Given the considerable variability among outcomes measures, we grouped them into anxiety and depressive symptoms. These included all such outcomes, whether measured by general or disorder-specific scales or subscales. As anxiety outcomes were sometimes measured for individuals without an anxiety disorder, we also conducted sensitivity analyses restricted to patients with one such disorder, diagnosed with a clinical interview or by use of a cut-off at a symptom scale. When a study used multiple measures from the same category, the average ES was computed using the CMA procedure (Borenstein et al., 2009) that assumes a correlation of 1 between outcomes. Since the correlation is probably less than 1, this approach is conservative (Scammacca et al., 2014). ITT data were preferred where available. If means and standard were not available, we calculated the SMD from other statistics available in the study, such as t-values or exact p-values, using the standard formulae in the program (Borenstein et al., 2009). If data was still insufficient for ES calculation, a request was sent to the study authors.

Drop-outs were defined as all randomized participants not finishing treatment, regardless of the reasons. Odds ratio (ORs) indicated the odds of participants dropping out from the VR versus the comparison group, with sub-unitary ORs indicating smaller odds for drop-out in the VR group.

We conducted separate meta-analyses for VR-enhanced therapy versus control, and respectively versus other active psychological treatments. Continuous outcomes (anxiety, depression) were pooled with a random effects model using the inverse-variance DerSimonian and Laird method (DerSimonian & Laird, 1986). For dichotomous outcomes, given we expected small trials, with some reporting few or no drop-outs, we used both the fixed effect Mantel-Haenszel method (Greenland & Robins, 1985; Mantel & Haenszel, 1959) with a continuity correction of 0.5 for zero counts, as well as Peto's method (Yusuf et al., 1985), as previously recommended (J. Cheng et al., 2016; J. P. T. Higgins & Green, 2011). Trials with zero drop-outs in both arms were excluded, due to concerns they might significantly inflate bias particularly in small trials (J. Cheng et al., 2016). We conducted sensitivity analyses excluding outliers and, respectively, excluding studies with a small number (N) of participants. Outliers were defined as studies in which the pooled ES's 95% CI was outside the 95% CI of the pooled ES (on both sides). We used an arbitrary cut-off of at least 25 randomized participants per arm for the analysis excluding small N studies. Though power calculations might differ from trial to trial, larger N trials are at least more precise in estimating the intervention effect (IntHout et al., 2015).

Heterogeneity was assessed with the I2 statistic, with values of 25%, 50% and respectively 75% indicating low, moderate and high heterogeneity (J. P. T. Higgins et al., 2003). We calculated 95% confidence intervals (CI) around I2 (J. P. A. Ioannidis et al., 2007), using the non-central χ2-based approach (Orsini et al., 2006). For categorical moderators, we conducted subgroup analyses using the mixed effects model, which uses a random-effects model within subgroups and a fixed-effects one across subgroups (Borenstein et al., 2009). For continuous moderators, meta-regression analyses employed a restricted maximum likelihood model with the Knapp-Hartung method (Borenstein et al., 2009).

We investigated small study effects and publication bias using a variety of methods. We resorted to visual inspection of the funnel plot, and contour enhanced funnel plots (Peters et al., 2008), where contour lines indicate regions where a test of treatment effects was significant for various established levels for statistical significance. We also employed statistical tests for small study effects. In the case of continuous outcomes, we conducted Egger's test (Egger et al., 1997) for the asymmetry of the funnel plot and corresponding Galbraith plots (Galbraith, 1988) if the test indicated significant asymmetry. We also used the trim and fill procedure (Duval & Tweedie, 2000) as a complementary method to adjust for potential publication bias or small study effects. For drop-out rates, as these were binary outcomes pooled with the ORs, we used the Harbord test (Harbord et al., 2006), which regresses Z/sqrt(V) against sqrt(V), where Z is the efficient score and V is Fisher's information (the variance of Z under the null hypothesis).

Data availability

The datasets generated and analysed during the current study are available in the Figshare repository, https://doi.org/10.6084/m9.figshare.5675407.

3.1.3. Results

Selection and inclusion of studies

The search generated 1394 records (720 after duplicate removal). We excluded 374 records based on abstract inspection and examined the full-texts for 346 articles. Figure 1 reports the flowchart of the inclusion process following the PRISMA guidelines (Moher et al., 2009). Subsequently, 42 trials met our inclusion criteria, six of which had insufficient data for ES calculation. Following contact with the original authors, we obtained data for one study. For two others, the author confirmed the samples overlapped with those from larger included studies. For 3 remaining trials, authors did not provide data, thus leaving a total of 39 trials in the meta-analysis (Supplementary Result).

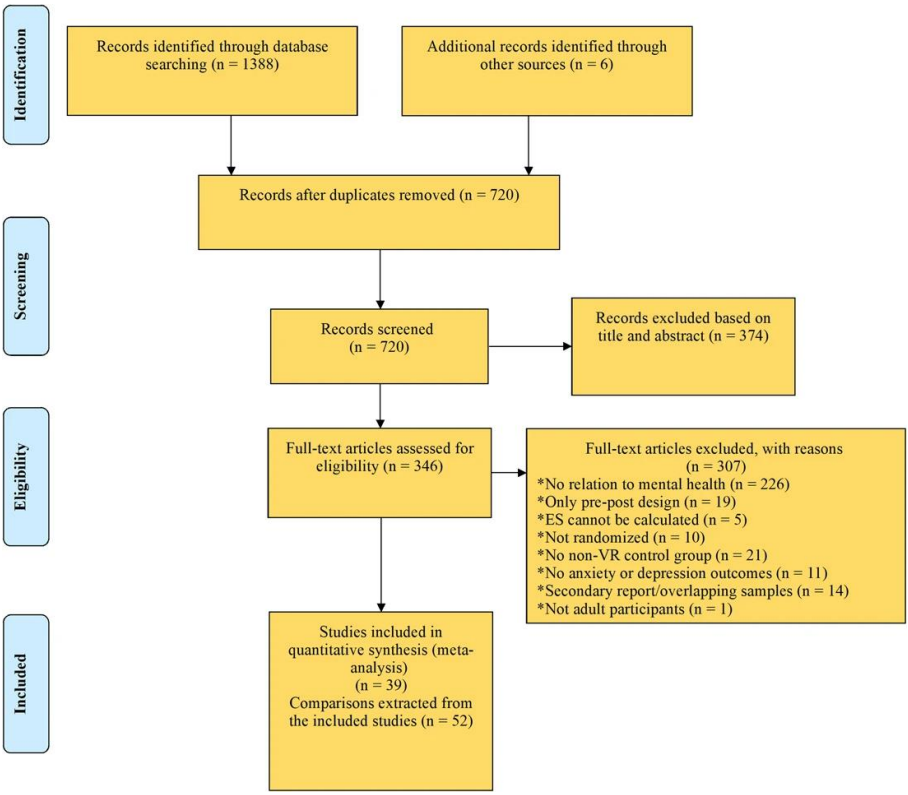

Figure 1. PRISMA flow-diagram of the study selection process

Characteristics of included studies

The 39 RCTs included 52 relevant comparisons, with 869 participants in the VR-enhanced condition, and 1122 in the control or active treatments ones. The most frequent conditions were anxiety and anxiety-related (e.g., PTSD) disorders (31 studies). The most frequently used VR therapy was VRE (in 21 out of the 39 RCTs), followed by VRCBT (in 19 out of the 39 RCTs). The number of VR sessions ranged from 1 to 16. The most used VR device was the head-mounted display (HMD) (35 studies). Apart from visual feedback, the majority of studies included sound (27 studies) or some form of navigation (18 studies). Only 6 trials explicitly assessed presence or immersion in the virtual environment. In most cases, developers of the VR program used were also among the authors (27 studies) (Table 2; Supplementary Table S1).

Risk of bias of the included studies

Most trials had uncertain or high risk of bias for three domains. Four RCTs had low RoB on all four domains. Nineteen studies were rated low RoB in only one

domain. For sequence generation and allocation concealment, the majority of trials (27 and respectively 28) did not provide any information to enable assessment. For blinding, only seven studies employed actual blinding of outcome assessors and 25 studies used exclusively self-report measures. For incomplete outcome data, 20 studies did not employ ITT analyses, and 9 studies did not include enough information to assess this domain. For this domain, we conducted additional subgroup analysis contrasting trials with low RoB versus the rest. Trials with high and unclear RoB were combined since given the ubiquity of treatment drop-out in RCTs, the lack of any mention of ITT strategies makes it very likely that none had been employed. For 3 trials, the number of drop-outs in one arm was unclear.

VR-enhanced therapy compared to a control condition

For anxiety outcomes twenty-three RCTs were pooled, g = 0.79, 95% CI 0.57 to 1.02, NNT = 2.36, with substantial heterogeneity (I^2 = 59%, 95% CI 35 to 74). Analyses restricted to participants with an anxiety disorder (17 comparisons) led to slightly smaller estimates: g = 0.72, 95% CI 0.51 to 0.94, NNT = 2.56, with similarly substantial heterogeneity (I^2 = 58%, 95% CI 28 to 76). Exclusion of three potential outliers led to a small decrease, g = 0.73, 95% CI 0.55 to 0.92, and reduced heterogeneity (I^2 = 36%; 95% CI 0 to 63). Only 7 trials had at least 25 participants randomized in each arm. Their aggregate ES was g = 0.64, 95% CI 0.39 to 0.88, and heterogeneity was still present (I^2 = 42%; 95% CI 0 to 76) (Table 3).

For depression, ten RCTs were pooled, g = 0.73, 95% CI 0.25 to 1.21, NNT = 2.54, with high heterogeneity (I^2 = 71%, 95% CI 45 to 85). Exclusion of one outlier resulted in a sizable decrease, g = 0.60, 95% CI 0.19 to 1.01, I^2 = 62%. Only one trial had at least 25 participants randomized in each arm.

Follow-up outcomes were only reported in two RCTs for anxiety and in one for depression.

Seventeen trials reported non-zero drop-outs in at least one group and nine trials reported zero drop-outs in both groups (Supplementary Table S2). Drop-out rates did not significantly differ between the groups, with similar estimations for the Mantel-Haenszel (OR = 1.34, 95% CI 0.95 to 1.89, $\chi 2 = 3.06$, p = 0.08) and Peto methods (OR = 1.37, 95% CI 0.96 to 1.95, $\chi 2 = 3.06$, p = 0.08).

VR-enhanced therapy compared to an active condition at post-treatment and follow-up

For anxiety twenty-nine RCTs were pooled, g = −0.02, 95% CI −0.14 to 0.10, with low heterogeneity (I^2 = 20%, 95% CI 0 to 50). Analyses restricted to trials with participants with an anxiety disorder (23 comparisons) also resulted in non-significant effects (albeit slightly more favourable to the non-VR interventions), g = −0.10, 95% CI −0.24 to 0.04, with similar heterogeneity estimates, I^2 = 26%, 95% CI 0 to 55. Results remained comparable after excluding two potential outliers, g = −0.02, 95% CI −0.13 to 0.08, I^2 = 0%, and in analyses limited to trials with at least 25 participants randomized per arm, g = −0.05, 95% CI −0.19 to 0.07, I^2 = 1% (Table 4).

For depression, thirteen RCTs were aggregated, g = 0.004, 95% CI: −0.20 to 0.21, with low heterogeneity (I^2 = 26%, 95% CI0 to 62). Exclusion of one outlier led to similar estimations, g = 0.07, 95% CI −0.10 to 0.25, I^2 = 0%, as did analyses excluding small N studies, g = −0.03, 95% CI −0.27 to 0.20, I^2 = 0%.

Follow-up anxiety outcomes were reported in 15 RCTs, g = −0.07, 95% CI −0.28 to 0.13, with moderate heterogeneity (I^2 = 40%, 95% CI 0 to 75). Results were similar with the exclusion of one outlier, g = −0.02, 95% CI −0.19 to 0.14, I2 = 8%. Depressive symptoms at follow-up were reported in 5 RCTs, g = −0.19, 95% CI −0.62 to 0.23, with moderate heterogeneity (I^2 = 57%).

Eighteen trials reported non-zero drop-outs in at least one group and ten trials reported zero drop-outs in both groups (Supplementary Table S2). Drop-out rates did not significantly differ between the groups, with similar results for the Mantel-Haenszel (OR = 1.05, 95% CI 0.77 to 1.43, $\chi 2 = 14.06$, p = 0.66) and Peto methods (OR = 1.05, 95% CI 0.77 to 1.43, $\chi 2 = 0.12$, p = 0.72).

Subgroup and meta-regression analyses

Recruitment setting was a significant moderator for the comparison between VR-enhanced interventions and control (p = 0.02) for anxiety, with the smallest ESs for recruitment from army settings and the highest for recruitment from a clinic. The type of anxiety disorder was also a significant moderator (p < 0.01), but this result is most likely affected by the high heterogeneity present within some of the small subgroups, as shown by the very large confidence intervals around I^2. Effects were very high for specific phobia (3 trials, g = 1.79, 95% CI 0.64 to 2.94) and panic disorder, though the latter was only studied in 2 trials. Effects were also high for flight anxiety (3 trials, g = 0.82, 95% CI 0.42 to 1.22). Effects were small for PTSD (4 trials, g = 0.39, 95% CI 0.04 to 0.74), and moderate for social anxiety (5 trials, g = 0.67, 95% CI 0.25 to 1.09). In the comparison with other active therapies, the type of VR intervention (VRE vs VR CBT) was a significant moderator (p = 0.02) for anxiety outcomes. In the subgroup (12 comparisons) where the VR-enhanced therapy was VRE, the non-VR intervention was slightly more effective (g = −0.18, 95% CI −0.35 to −0.006). In this subgroup, the non-VR intervention consisted of imaginal exposure (6 comparisons), CBT (2 comparisons) and in vivo exposure (4 comparisons) (Tables 3 and 4).

Univariate meta-regression indicated significant negative relationships between publication year and both anxiety (slope = −0.06, 95% CI: −0.09 to −0.03) and depression ESs (slope = −0.10, 95% CI: −0.18 to −0.02) in comparison with control conditions, which were maintained in sensitivity analyses excluding outliers. The number of elements of interaction with the virtual environment was positively associated with anxiety outcomes (slope = 0.22, 95% CI: 0.01 to 0.42), but this result did not survive in a sensitivity analysis excluding outliers. For the contrast with other active conditions, publication year, mean age and respectively RoB score were significantly related to anxiety ESs, but only the relationship with age (slope = 0.02, 95% CI: 0.006 to 0.04) survived in analyses excluding outliers.

Small study effects and publication bias

Visual inspection pointed to an asymmetrical funnel for both anxiety and depression. Contour enhanced funnel plots showed that for anxiety most of the studies with higher standard errors had results overcoming conventional statistical threshold of p < 0.05, with a considerable proportion of these even significant at the more conservative threshold of p < 0.01. Results were similar for depression though the number of ESs was much smaller. Egger's regression intercept test was statistically significant for both anxiety (intercept = 2.03, 95% CI 0.07 to 3.98, p = 0.04) and

depression outcomes (intercept = 3.24, 95% CI 0.10 to 6.39, p = 0.04). Galbraith plots for anxiety evidenced the same pattern, as studies with low precision (i.e., inverse of the standard error) did not scatter randomly around the regression line, with most of them having effect estimations benefiting the VR intervention. For depression the pattern was inconclusive, probably due to the small number of studies. Finally, Duval and Tweedie's trim and fill procedure also pointed to small study effects for anxiety and depression. For anxiety, adjustment for potentially missing studies (n = 5), was associated with the ES decreasing from 0.79 to 0.62, whereas for depression (n = 3), it rendered the pooled ES non-significant. There was reduced indication of small study effects or publication bias for the comparison with other active treatments, with Egger's test non-significant and no adjustment for missing studies, except for depression.

For drop-out rates, the Harbord test did not indicate small study effects (coeff = 0.16, 95% CI −1.92 to 2.24, p = 0.87). However, it is important to note this analysis may be biased, as it excluded studies with zero drop-out counts in both arms, which were also some of the smaller N studies (Supplementary Table S2).

3.1.4. Discussion and conclusions

In the reported meta-analysis, we showed moderate to large effects of VR interventions compared to control conditions (e.g., waitlist, placebo, relaxation, treatment as usual), for anxiety and depression outcomes. The number of studies with follow-up evaluations was too small for a meaningful ES estimation. There was moderate to high heterogeneity and a number of studies with extreme values. Most studies had a small number of participants and there was substantial evidence of small study effects for anxiety outcomes, pointing to potential publication bias. The limited number of studies reporting on depression outcomes precluded us from drawing a meaningful conclusion about small study effects. Adjustment for funnel plot asymmetry, as well as sensitivity analyses excluding outliers or restricted to studies with a moderate number of randomized participants per arm reduced the pooled ES for anxiety, though it still remained moderate to large. Only 7 trials that reported on anxiety outcomes had randomized at least 25 participants in each arm. The persistent evidence of small study effects, as well as the significant heterogeneity, casts doubts over the reliability of the large effects observed for anxiety (Dechartres et al., 2014; J. P. A. Ioannidis et al., 2007; Nüesch et al., 2010). Heterogeneity continued to remain moderate with large confidence intervals even when extreme values were excluded, showing it was not simply the by-product of a few trials. Two thirds of the studies used waitlist controls, and effect sizes were large in waitlist comparisons. Use of waitlist controls might inadvertently and artificially inflate effect sizes for both anxiety and depression outcomes (Cuijpers et al., 2016; Furukawa et al., 2014).

Conversely, compared with established active interventions, effect sizes were non-significant for both anxiety and depression outcomes, at post-test and follow-up. Heterogeneity was small to moderate and there was limited evidence of funnel plot asymmetry or small study effects. Sensitivity analyses excluding outliers or restricted to studies with at least 25 participants randomized in each arm produced similar estimations. There were more trials in the latter category (12) than in the comparison with control conditions (7), but these were still a minority. All but one of the trials were powered to test superiority, not equivalence or non-inferiority (Christensen, 2007), so it would be premature to construe our findings as proof of equivalent effects. Most frequently employed non-VR active interventions were IVE and CBT, both shown to be effective for anxiety and depression, thereby potentially difficult to outperform.

VR-enhanced interventions did not improve attrition, producing similar drop-out rates with control conditions and other active interventions. These findings contradict previous speculation of possible comparative benefit (Botella et al., 2015; Freeman et al., 2017; Meyerbröker & Emmelkamp, 2010). However, most trials were small and many reported zero drop-outs, sometimes in both arms, so the stability of this result needs to be considered with caution. We were not able to evidence small study effects for analyses on attrition, but this result is most likely biased by the fact studies with zero counts in both arms were excluded and many of these were also small studies.

The vast majority of RCTs of VR interventions had high or uncertain risk of bias across domains. Two previous meta-analyses (McCann et al., 2014; Turner & Casey, 2014) examined bias using combinations of instruments, which included aspects not linked to any type of trial bias (e.g., training for providers), potentially obfuscating distorting effects. In contrast, we used the Cochrane Risk of Bias tool (J. P. T. Higgins et al., 2011), which evaluates domains likely to distort outcomes. Only four trials could be rated as low RoB on all domains considered, preventing us from reliably assessing the relationship between overall trial risk of bias and outcomes. The only RoB domain where most trials reported information was incomplete outcome data. Almost two thirds of the studies were rated as high risk of attrition bias, again questioning the reliability of the ES estimations, as exclusion of participants from RCT analyses was shown to distort outcomes (Abraha et al., 2015; Nüesch et al., 2009). In exploratory subgroup analysis, we did not find differences between studies with high/uncertain versus low RoB for incomplete outcome reporting, though the number of studies with low RoB was small (seven), particularly in comparisons with control. It is possible previous assessments concluding no relationship between trial risk of bias and ESs might have been too optimistic.

Though the presence of the developers of VR interventions among the author pool was not significantly associated with changes in the magnitude of the effects, it is worth underscoring the vast majority of trials did involve such a developer. For instance, for the comparison with control conditions, only five anxiety effect sizes came from independent studies, and 17 from trials involving the developer. As such, it is possible that the insufficient variability in our sample of included trials prevented us from detecting more subtle differences. Moreover, we only examined whether one of the authors had also developed the VR treatment program used, not any potential commercial involvements with VR companies, which could arguably represent a more direct conflict of interest. However, since most articles did not report this information, we could not examine it systematically.

We identified few moderators, owing to the fact most subgroups were small and affected by high heterogeneity within the group. Recruitment setting seemed to have an influence on ESs in comparisons between VR-enhanced and control conditions, with smaller effects for recruitment from army settings, but this may also be a spurious result since some of the subgroups contained a very limited number of studies. Type of anxiety diagnosis also appeared to be a significant moderator, with high effects for specific phobia and flight anxiety, and moderate or small effects for social anxiety and PTSD. It is likely that this is at least partly a spurious result, given subgroups were small and heterogeneity was high in all of them. The type of active comparison intervention used appeared to matter, with VR-enhanced exposure having slightly smaller effects than non-VR interventions. Again, the number of studies was

small and this relationship could have also been confounded by other variables, such as the type of problem for which the therapy was used.

It was speculated (Freeman et al., 2017) that improved engagement with the virtual environment, as measured by immersion or a sense of presence, could play an important role in the effectiveness of VR. Only a modest number of trials measured immersion and presence explicitly. Even in those that did, most did not analyze these variables in relationship to treatment outcomes or found no association. We showed that the number of elements employed by the VR technology, a crude indicator of interaction, was positively related with anxiety outcomes in comparisons with passive, but not active treatments. However, this result did not survive sensitivity analyses and could be an artifactual finding. But even for visual stimulation, though one might assume that more recent studies use very sophisticated technology, instead of stereoscopic simulations not intended for VR use, we saw no evidence to this effect. For example a 2017 trial (Bouchard et al., 2017) relied on the same technology as similar trials from 2013 (Anderson et al., 2013) and even 2005 (Klinger et al., 2005).

Publication year was consistently negatively associated with outcomes, though reasons for this trend remained unclear. A rise in larger or lower risk of bias trials seems unlikely given we observed few such trials. The apparent decrease in effectiveness with the passing of time might also be a by-product of the early use of pilot, low powered studies where only large effects can overcome the significance threshold, a strong initial publication bias for positive findings, as well as time lag bias, whereby studies with positive results are published first and dominate the field, until the negative, but equally important, studies get published (J. P. T. Higgins & Green, 2011; J. P. Ioannidis, 1998). Previous meta-analyses of RCTs of VR interventions either did not consider publication bias at all (McCann et al., 2014; Opriş et al., 2012), or reported optimistic estimates (Turner & Casey, 2014), based on the fail-safe N, whose use is discouraged for being unreliable and misleading (J. P. T. Higgins & Green, 2011). We used a range of methods to assess funnel plot asymmetry, all of which corroborated that small studies were numerous, mostly significant and overestimated effects for comparisons with control conditions. Publication bias for positive findings, probably more prominent in the early years of studying VR interventions, is one likely cause of small study effects. We conjecture it is most likely present in the literature of VR interventions for anxiety, where most trials are concentrated.

There are several limitations to our meta-analysis. There was a high degree of heterogeneity, particularly in comparisons with control conditions. This was accompanied by very large confidence intervals around I2, even for the comparisons where heterogeneity estimates were smaller. Residual heterogeneity persisted even after sensitivity analyses were conducted, or potential moderators explored. NNTs can be useful as an ancillary clinical ES measure, but there is disagreement regarding the most adequate calculation method (Furukawa & Leucht, 2011), and concerns over their potential to mislead, particularly when resulting from meta-analyses, as baseline risk can vary substantially between trials (Smeeth et al., 1999). Many of the subgroup analyses were underpowered and we were able to identify few moderators. We could not calculate effect sizes for three trials where the report did not contain enough information and the original authors did not provide the data. However, given their size and the total number of included trials, their exclusion is unlikely to have influenced estimations.

From the standpoint of dissemination and implementation, our results leave several open questions. Virtual reality enhanced interventions had moderate to large effects compared to control conditions, though these effects were likely inflated by several factors in the design and implementation of the trials. We could find few differences with other active interventions. These might be construed as evidence VR-enhanced interventions could be added to the armamentarium, as another effective choice available to clinicians and patients.

However, other key aspects remain unclear. Though it would be intuitive to consider VR-enhanced interventions as more cost-effective than traditional anxiety treatments, notably in vivo exposure, research substantiating this claim is missing. Moreover, it might hinge on the specific disorder targeted. For instance, for flight anxiety it may seem evident that it would be more cost-effective to conduct VR-enhanced exposure than buy a plane ticket for in vivo exposure. Conversely, for height anxiety, it could be more cost-effective to scale a flight of stairs with a patient, than to purchase a HMD system and pay for the software development of a fully immersive VR environment. Nonetheless, this kind of tailored, immersive and sophisticated technology does not seem to be used much, even in recent trials, further complicating a realistic calculation of cost-effectiveness. One might also argue VR-enhanced interventions might be particularly suitable for disorders where other active interventions have been less effective. Nonetheless, in the case of one such disorder-post-traumatic stress disorder- two recent trials (McLay et al., 2017; Reger et al., 2016) failed to find additional benefits for VR interventions over non-VR treatments such as prolonged exposure, both in terms of primary outcome, as well as drop-out rates, with follow-up results actually better for the non-VR intervention.

Most importantly, many existent trials are poorly reported and exposed to bias. The effort to move forward should primarily focus on elevating the quality of VR trials. Larger trials minimizing risk of bias by prospective registration and transparent and complete reporting, as well as using credible control groups, are necessary. A recent ongoing trial described in a published protocol is one such example (Miloff et al., 2016). Trials should also report cost-effectiveness analyses in an attempt to clarify whether and under which conditions are VR-enhanced treatments cost-effective. Finally, they should include an evaluation of the participants' engagement with the VR environment, so as to clarify how immersive and sophisticated the system needs to be to support improved outcomes. Moreover, given the predominance of trials conducted by developers of VR treatments, independently conducted trials are also critical. It is essential that negative results are afforded journal space in order to tackle potential publication bias.

3.2. Study 2: The efficacy of cognitive bias modification interventions in anxiety and depressive disorders: a network meta-analysis[2]

3.2.1. Introduction

The current research agenda for psychological treatments recommends moving towards developing interventions mechanistically, by translating experimental findings (Holmes et al., 2018). Cognitive bias modification (CBM) interventions are prototypical examples. These encompass a diversity of approaches with multiple variants in each, such as attention bias modification (ABM), interpretation bias modification (CBMI), or approach and avoidance training (AAT). Across all, a target cognitive bias is manipulated, with participants taught, often without being explicitly made aware, to preferentially attend to, process or otherwise engage with certain types of stimuli (i.e., positive, neutral), while simultaneously avoiding others (i.e., negative, threatening) (MacLeod & Mathews, 2012). CBM interventions are appealing due to accessibility and scalability, as they consist of brief sessions of a computer-based task, possibly administered online.

Nevertheless, the effectiveness of CBM interventions is contentious. Meta-analyses of randomized controlled trials (RCTs) often reached strikingly different conclusions. For the most investigated form (i.e., ABM for anxiety disorders) some meta-analyses reported small, frequently non-significant, symptom reductions compared to control conditions (Cristea et al., 2015; Heeren et al., 2015), while others reported significant effects of larger magnitude (Linetzky et al., 2015; Price et al., 2016). There were fewer trials for depression, with mixed findings (Cristea et al., 2015; Menne-Lothmann et al., 2014).

The diversity of CBM procedures is mirrored by a variety of control groups, rendering the standard for gauging the effectiveness of CBM uncertain. For instance, owing to the computerized tasks, usually not requiring participant awareness, several studies have employed a control condition aimed to function as a "placebo". This "no contingency" or "sham training" control task is identical to the active intervention, without favoring a stimulus type, i.e., positive or neutral stimuli appear as frequently as negative ones. As with placebo, some studies reported benefits for participants randomized to this control arm (Boettcher et al., 2013). Furthermore, it is unclear whether certain versions of CBM are more effective than others for specific symptoms.

Due to the scarcity of studies comparing strains of CBM among each other and with different control groups, these questions cannot be settled in a typical pairwise meta-analysis of direct comparisons. Conversely, network meta-analyses (NMAs) synthesize direct and indirect evidence enabling the estimation of comparative effects even in the absence of trials directly comparing interventions (Leucht et al., 2016).

We therefore conducted NMAs to determine the relative effectiveness of CBM procedures (i.e., ABM, CBMI, AAT), compared among each other and with control groups, for anxious and depressive symptomatology.

---

[2] This study has been published
Fodor, L. A., Georgescu, R., Cuijpers, P., Szamoskozi, Ş., David, D., Furukawa, T. A., & Cristea, I. A. (2020). Efficacy of cognitive bias modification interventions in anxiety and depressive disorders: A systematic review and network meta-analysis. The Lancet Psychiatry, 7(6), 506–514. https://doi.org/10.1016/S2215-0366(20)30130-9

3.2.2. Methods

Data availability statement, protocol registration and reporting guidelines

The meta-analysis was prospectively registered (PROSPERO registration CRD42018086113) and reported following the PRISMA extension for Network Meta-Analyses (Hutton et al., 2015).

Identification and selection of studies

A literature search in PubMed, PsycINFO, EMBASE and Cochrane Central Register of Controlled Trials databases was conducted through February 7th 2020, using the combinations of terms (both as controlled vocabulary thesaurus and free-text) relating to "cognitive bias modification", "attention* bias modification", "attention* bias training", ", "interpret* bias modification", and "anx*", "fear", "depress*", "dysth*", "obsess*", "phob*", "panic", "agoraphob*", "ptsd", "post traumatic", "acute stress", "adjustment disorder" (see the Appendix for complete search strings). We also inspected references from the most recent systematic reviews and meta-analyses (Cristea et al., 2015; Hallion & Ruscio, 2011; Heeren et al., 2015; Liu et al., 2017). Peer-review publications in English, Romanian, Spanish, Italian, German and Dutch were considered.

Eligible studies were RCTs comparing a CBM intervention to a control condition for anxious or depressive symptom outcomes measured on validated clinical scales, in adults whose primary complaint consisted of symptoms of anxiety or depression, either diagnosed, with a diagnostic interview (e.g., Structured Clinical Interview for DSM-IV) or a validated clinical scale (e.g., Liebowitz Social Anxiety Scale/LSAS), or of subclinical intensity evaluated on a validated clinical scale. Participants with comorbid anxious or depressive symptoms were eligible. Disorders were defined according to the DSM-IV/IV-TR, as recruitment in most trials likely predated the DSM-5. Combination studies of CBM and another intervention were eligible, provided the control group also received the ancillary intervention. Studies contrasting CBM with non-CBM active intervention (e.g., cognitive behavior therapy) were excluded. State measures of anxiety or depressed mood were ineligible because they do not reliably index symptoms of clinical importance.

Two researchers (LAF and RG) independently screened all abstracts, subsequently examined full-texts and selected eligible RCTs. All disagreements were resolved by discussion and consultation with a third author (IAC) until consensus was reached.

Data extraction

We extracted information about: (1) Sample: clinical (diagnosed) or subclinical (elevated symptoms); (2) Total number of participants randomized (N); (3) CBM intervention: AAT; ABM; CBM-I (definitions in Table 1); (4) Control condition: Sham training (SHAM), Opposite ABM (OABM); Waitlist (WL) (Table 1); (5) Number of CBM sessions; (6) Delivery: laboratory, home, clinic or combinations; 7) Outcome measures for anxiety and depression; and 8) Publication year.

Primary outcomes

We expected trials to employ multiple outcome measures, hence we pre-specified a hierarchy. For studies reporting both anxiety and depression outcomes, we first considered the investigator-declared primary outcome. If none was identified, we selected it based on the focus of the intervention, e.g., anxiety outcomes for anxiety disorders. Clinician-based instruments were favored over self-report, if available.

Secondary outcomes

As anxiety and depression are highly comorbid (Lamers et al., 2011), we also considered comorbid depression (i.e., in trials of CBM for anxiety disorders) and anxiety (i.e., in trials of CBM interventions for depressive disorders) outcomes.

Risk of bias

We used the Risk of Bias assessment tool, developed by the Cochrane Collaboration (J. P. Higgins et al., 2016), which assesses possible sources of bias in RCTs. The following domains were rated: a) random sequence generation, b) allocation concealment, c) blinding of participants and personnel, d) blinding of outcome assessors, e) incomplete outcome data and f) selective outcome reporting. Domain c) was considered low risk if blinding of participants was attempted, regardless of whether subsequent checks were performed to determine if it was maintained. For domain d), clinician-based measures were prioritized. For self-report, participants were considered their own assessors (J. P. Higgins et al., 2016), with ratings of low risk given if they were blinded to the intervention received. Domain e) was assessed as low risk if all randomized participants were included in the analysis, through the use of an intent-to-treat (ITT) approach or complete data availability. Domain f) was assessed as low risk if primary and secondary outcomes were pre-specified in a prospectively registered protocol or trial registration, with no substantial changes between registration and publication. Retrospectively or non-registered studies were rated as unclear.

Two independent researchers (LAF, RG) extracted outcome data and rated risk of bias, with disagreements resolved by consensus after discussion with another author (IAC).

Meta-analysis

All analyses were conducted in STATA/SE 15 (StataCorp.2017, 2017) (the "network" and "mvmeta" packages (Chaimani et al., 2013; I. White, 2015; I. R. White, 2011)) and R (R Core, 2018) (package "netmeta" (Rucker et al., 2019)).

Means, standard deviations (SD) and sample sizes in each arm were used to calculate between-groups effect sizes (ES) as post-intervention standardized mean differences (SMD) and corresponding 95% confidence interval (CI). The SMD represents the difference in means between the intervention and control arms divided by the pooled standard deviation. ITT data were preferred, when available. For studies with two or more arms from the same category (i.e., a type of CBM or of control), we extracted data from those most similar to the standard version. For example, if a study contained both CBM-I and a modified version of it with additional components, we chose the former. If data were insufficient for ES calculation, study authors were contacted. We employed multivariate random effects meta-analysis with restricted maximum likelihood (REML) estimator to conduct four NMAs (one per outcome). We graphically represented results as network plots, whereby the size of the nodes is directly proportional with the number of patients, while the thickness of the lines connecting the nodes is weighted by the number of trials directly assessing the comparison. Additionally, we constructed network plots that incorporated risk of bias for each domain rated, using colored edges to represent low, high and unclear risk of bias. The comparison-specific bias level was set as the rating in the majority of studies in each comparison (i.e., the mode) (Chaimani et al., 2013).

The transitivity assumption was evaluated by visually inspecting relevant study characteristics. Based on previous literature, we considered two potential effect modifiers (number of sessions and delivery setting) and examined their distributions

across comparisons. Network consistency, the extent to which included studies are comparable, both statistically and substantively (J. Higgins et al., 2012), was evaluated with three methods. First, to detect significant overall inconsistency, we used a design-by-treatment interaction model with a global Wald statistic which under consistency follows a $\chi 2$ distribution (Donegan et al., 2013; I. R. White et al., 2012) (non-significant p values indicate no inconsistency). Second, we used a loop-specific approach to estimate the inconsistency factor (IF) in each loop as the absolute difference between direct and indirect estimates and truncating the confidence intervals to zero, using a Z-test to decide if inconsistency is significant (Veroniki et al., 2013) (i.e., the lower limit of the IF's 95% CI touches zero). Thirdly, we employed a side-splitting method, a frequentist adaptation of the original hierarchical Bayesian method (Dias et al., 2010; Donegan et al., 2013). It reports the estimated direct and indirect treatment effects and their difference, with consistency inferred based on the p-value for the difference.
Contribution plots displayed the differential contributions of direct comparisons to the network summary effect. Interventions were ranked by calculating the surface under the cumulative ranking (SUCRA), which denotes the probability (in percentages) of superior effectiveness for each intervention compared to a theoretical ideal (i.e., always the best without uncertainty) intervention.

Heterogeneity was investigated by displaying forest plots, including summary effects along with their 95% CI and their corresponding 95% prediction intervals (PrI's) for all comparisons. Prediction intervals represent confidence intervals of the approximate predictive distribution of future trials, considering heterogeneity (J. P. T. Higgins et al., 2009). We further conducted three sensitivity analyses excluding studies: (1) employing AAT, initially devised for addiction (Cristea et al., 2016); (2) on PTSD participants, where better outcomes were reported for SHAM than for ABM (Badura-Brack et al., 2015); (3) excluding studies where participants in the SHAM intervention were not exposed to any contingency (e.g., neutral scenarios). We employed network restricted maximum likelihood meta-regression (I. R. White, 2011) using "mvmeta" to examine two possible moderators for the comparison between CBM and SHAM - number of treatment sessions (continuous) and delivery setting (recoded dichotomously as laboratory versus others).

Small study effects were examined through visualization of comparison-adjusted funnel plots and with Egger's linear regression test of funnel plot asymmetry (Egger et al., 1997). Interventions were ordered such that all active interventions were contrasted sequentially to WL, SHAM and OABM control conditions (Chaimani et al., 2013).

3.2.3. Results

The search generated 2125 records (1156 after duplicate removal). We excluded 854 records based on abstract inspection and examined 302 full-texts. The PRISMA flowchart (Moher et al., 2009) (Figure 1) reports the inclusion process. We contacted authors of 8 studies with insufficient ES data and retrieved datasets for 1. Consequently, 82 reports describing 85 separate trials were included in the NMA.

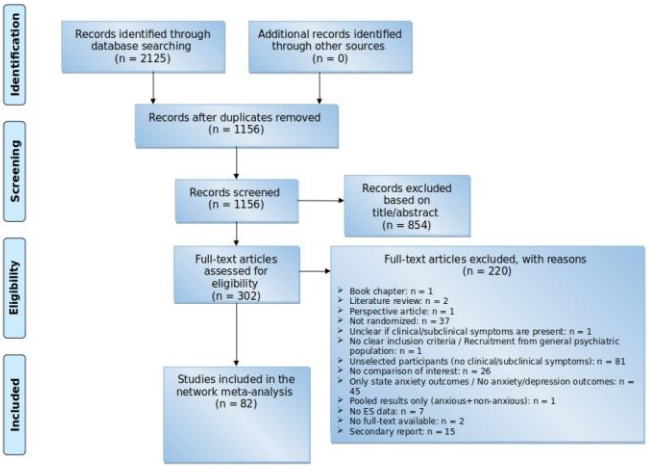

Figure 1. PRISMA flow-diagram of the study selection process

Sixty-five trials (2026 treated and 1871 control participants) focused on anxiety-related disorders, of clinical or subclinical intensity. Twenty trials (544 treated and 572 control participants) focused on depressive disorders or symptoms. Delivery settings included the laboratory (44 studies), online/at home (21), clinic/hospital (8) or a combination of these (12). Treatment sessions ranged from 1 to 84, with 12 RCTs employing one, and 43 RCTs 8 sessions or more. Most trials employed sham training and directly compared ABM and SHAM (Figure 2). Comorbid depression outcomes were reported in 31 studies (1101 treated, 1070 control participants) on anxiety disorders. Conversely, comorbid anxiety outcomes were present in 11 studies on depressive disorders (250 treated, 251 control).

Most trials had uncertain or high risk of bias for five out of six domains. Four RCTs had low risk of bias for all domains, while six RCTs had low risk for five domains. Sequence generation was rated low risk in 31 trials (42 had insufficient information), allocation concealment in 13 trials (67 insufficient information) and blinding of participants and personnel in 37 trials (38 unclear). Blinding of outcome assessors was rated low risk in 39 trials (69 used self-report measures exclusively). For incomplete outcome data, 54 studies reported ITT analyses or complete outcome data were available. For selective outcome reporting, 12 studies were rated low risk.

For the primary outcome of anxiety, the network plot (Figure 2a) showed a well-connected network, consisting of 8 nodes. The majority of direct comparisons were at unclear risk of bias, except for blinding of outcome assessment (unclear/high) and incomplete outcome data (low). Across methods, there was no evidence for inconsistency. In the NMA only CBMI significantly reduced the anxiety compared to WL (SMD = -0.55, 95% CI: -0.91 to -0.19) or SHAM (SMD = -0.30, 95% CI: -0.50 to -0.10). However, prediction intervals for these comparisons were large and included 0. SUCRA probabilities indicated that AAT and CBMI presented the greatest likelihood

of reducing anxiety outcomes (both ~77%). Egger's test did not detect funnel plot asymmetry, t(80) = 0.31, p = 0.757.

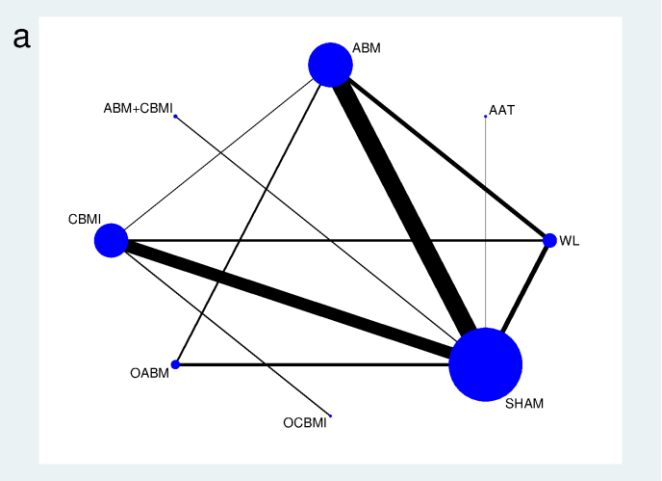

Figure 2a. Network geometry of the CBM interventions for anxiety outcomes

For the primary depression outcome, the network plot (Figure 2b) shows a well-connected network of CBM interventions, except for ABM+CBMI. Across all methods, there was evidence for inconsistency. Risk of bias, the contribution plot and SUCRA probabilities are presented in the Appendix. In the NMA CBMI significantly reduced depression compared to WL (SMD = -0.63, 95% CI: -1.04 to -0.23). The 95% prediction interval was large and included 0. Other statistically significant differences involved the singly-connected ABM+CBMI node. Egger's test detected funnel plot asymmetry, t(22) = -2.10, p = 0.047.

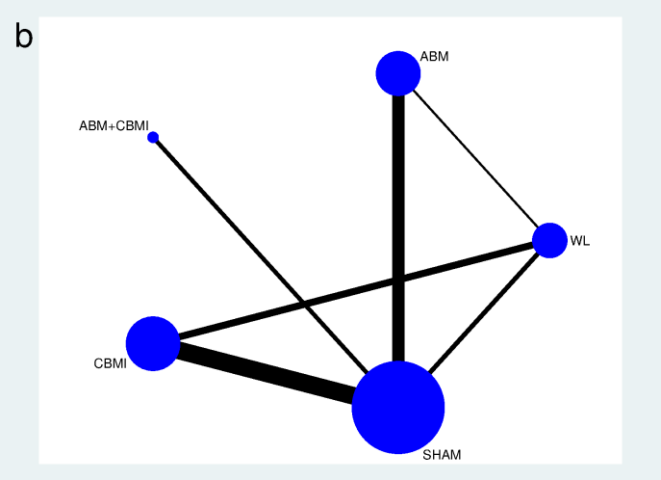

Figure 2b. Network geometry of the CBM interventions for depression outcomes

For the secondary outcome of comorbid depression (in anxiety RCTs), the network plot showed a well-connected network of CBM interventions, consisting of 6 nodes. None of the methods revealed evidence of inconsistency. Risk of bias, the contribution plot and SUCRA probabilities are displayed in the Appendix. In the NMA, only CBMI significantly reduced depression compared to WL (SMD = -0.42, 95% CI: -0.68 to -0.15), SHAM (SMD = -0.21, 95% CI: -0.41 to -0.01) and ABM (SMD = -0.24, 95% CI: -0.46 to -0.01). However, all 95% prediction intervals were large and, except for CBMI versus WL, included 0. Egger's test did not detect funnel plot asymmetry, $t(42) = -1.34$, $p = 0.187$.

For the secondary outcome of comorbid anxiety in depression RCTs, the network plot evidenced a well-connected network of CBM interventions, consisting of 4 nodes. The majority of the comparisons were at unclear and high risk of bias. Evidence for inconsistency was mixed, with 2 of the 3 methods indicating inconsistency. The NMA showed no significant differences and Egger's test did not detect funnel plot asymmetry, $t(11) = 0.74$, $p = 0.472$.

Post-hoc sensitivity analyses closely replicated the main analyses, with a few differences. With the exclusion of PTSD trials (n=7), ABM significantly reduced anxiety compared to WL (SMD = -0.35, 95% CI: -0.59 to -0.12) and SHAM (SMD = -0.16, 95% CI: -0.28 to -0.04). Excluding trials where SHAM participants were not exposed to any contingency (n= 15), CBMI significantly reduced anxiety compared to WL (SMD = -0.62, 95% CI: -1.07 to -0.18), but not SHAM.

Meta-regression analyses showed that the number of treatment sessions was not significantly related to outcomes. Delivery setting was a significant moderator only for the ABM versus SHAM comparison (beta = 0.44. 95% CI 0.10 to 0.77) for anxiety outcomes.

3.2.4. Discussion and conclusions

In a network meta-analysis of 85 trials, CBM interventions showed limited benefits over control conditions, for both anxious and depressive symptomatology. In 65 trials on anxious participants, CBMI outperformed waitlist (SMD = -0.55, 95% CI: -0.91 to -0.19) or sham training (SMD = -0.30, 95% CI: -0.50 to -0.10) for anxiety outcomes. However, prediction intervals were large and contained the SMD of 0, suggesting that the effects of future similar trials could fluctuate across a wide range of effects. Similar results were reported for comorbid depression outcomes, present in around half of the trials, suggesting that the effects of CBMI might be disorder- rather than symptom-specific. These effects are modest compared to similarly delivered internet-based cognitive behavioral interventions for anxiety disorders (SMDs compared to mostly waitlist control ranging from 0.70 for generalized anxiety disorder to 1.31 for panic disorders) (Andrews et al., 2018). In post-hoc analyses excluding the more inert type of SHAM (neutral scenarios), only differences between CBMI and waitlist remained significant. Few differences emerged among CBM interventions, except for the superiority of CBMI over ABM for comorbid depression.

For ABM, the only significant findings consisted of small effects compared to waitlist and sham on primary anxiety outcomes, in sensitivity analyses excluding PTSD trials. Our definition of anxiety disorders predated the DSM-5, hence including stress-related disorders (all included PTSD trials, relied on the DSM-IV-TR). More generally, findings for ABM corroborate previous meta-analyses reporting very similar estimates (Cristea et al., 2015; Heeren et al., 2015), but contradict others reporting larger effects of ABM for participants with clinical anxiety (Linetzky et al., 2015; Price et al., 2016). A pairwise meta-analysis of 11 studies (Linetzky et al., 2015) showed moderate effects of ABM for clinician-rated, SMD=0.42, 95% CI 0.18 to 0.66, but not self-reported anxiety. In an individual participant data meta-analysis of 13 trials (Price et al., 2016), the authors reported significant effects of ABM on diagnostic remission (OR= 2.57, 95% CI 1.31 to 5.22), but not on the continuous measure, the clinician-administered LSAS. Laboratory delivery was associated with better outcomes for ABM versus SHAM, corroborating previous reports (Cristea et al., 2015; Heeren et al., 2015).

In the considerably fewer trials on depressed participants (n = 20), only CBMI outperformed waitlist for primary depression outcomes (SMD = -0.57, 95% CI: -0.99 to -0.16). However, due to evidence of network inconsistency, this finding might be unreliable. The network geometry revealed one open, singly-connected, node for the ABM+CBMI combination. Aside from a direct comparison with SHAM, the whole evidence for the combined intervention was indirect, rendering the very large effects observed not credible. Hence, the effectiveness of the combined treatment cannot be established. We used a hierarchy of outcomes, favoring the investigator-declared primary outcomes and, barring that, clinician-based over self-report measures. Therefore, it is unlikely that our findings are explained by the choice of measures. Importantly, anxiety trials formed well-connected networks for both anxiety and depression outcomes, with no evidence for inconsistency, making fundamental differences between trials unlikely and further supporting the robustness of the findings.

Few differences emerged among the various control conditions employed. Waitlist was always nominally inferior to sham CBM, supporting the notion that interventions should be compared with more adequate, active control groups (Cristea, 2019). By contrast, a strength of CBM trials is the frequent inclusion of a sham

condition, in which participants are not encouraged to preferentially process a certain stimulus type. Analogous to pill placebo, these active control conditions can be targeted (50:50 ratio of targeted versus non-targeted emotional stimuli) or neutral (neutral stimuli) Attempts to blind participants to their group allocations add to the placebo similarity, a rare occurrence in research on psychological interventions. Yet just under half of the trials specifically mentioned participant blinding, with an almost equal number lacking information. Therefore, it is difficult to ascertain whether the sham condition truly remained as such. Interestingly, in post-hoc sensitivity analyses excluding studies using the neutral version of sham as control, CBMI was only superior to waitlist. Finally, although sham training does not ostensibly encourage preferential processing, it might nonetheless have beneficial emotional effects (Blackwell et al., 2017) by repeatedly exposing participants to bias contingencies and enhancing, for example, psychological flexibility (Edwards et al., 2018). Nonetheless, pill placebo is also often not inert (Blackwell et al., 2017). Furthermore, some trials included a seemingly paradoxical "opposite" intervention (i.e., OABM), in which contingencies are modified so that negative stimuli are preferentially attended to or processed, in an effort to increase cognitive bias. If the relationship between attention bias and anxiety symptoms is, as hypothesized (MacLeod & Mathews, 2012), causal, this control arm could be expected to be deleterious, similarly to a nocebo. Nonetheless, OABM did not diverge from other control conditions, though, owing to ethical reasons, few trials used it.

Our network meta-analysis is not without limitations. For most trials, owing to insufficient information, risk of bias was rated as uncertain across several of the domains considered. The minuscule number of studies with low overall risk precluded us from running further sensitivity analyses. The conduct and reporting of trials of CBM interventions is still wanting, as previously shown (Cristea et al., 2015). Consequently, as high or unclear risk of bias were associated with exaggerated effect estimates across interventions (Savović et al., 2018), the effects reported might overestimate the "true" effects. Other methods of assessing evidential value (Amad et al., 2019; Sakaluk et al., 2019) might complement a more in-depth assessment of trials with uncertain risk. We found little statistical evidence of small study effects in any network. However, publication bias cannot be completely ruled out without directly identifying unpublished trials. We excluded the trials in the grey literature like dissertations and conference proceedings, owing to concerns about insufficient data reporting and reliably connecting these reports with journal articles. For depressed participants, trials were few and there was evidence of inconsistency in the resultant networks, meaning findings might be unreliable.

As commonly seen in psychological treatment research, clinical heterogeneity is unavoidable. Both ABM and CBMI denote families of approaches targeting a particular process, with variations in the tasks and stimuli employed, use of additional components and doses. Clinical heterogeneity extends to control conditions, particularly sham training, though most studies used the no contingency (i.e., 50:50 ratio) arm. For the purposes of the NMA, sham training conditions were also considered interchangeable across CBM interventions. Though their principle is the same (i.e., participants are not encouraged to preferentially process a certain stimulus type), there is variation in the nature of the stimuli (e.g., faces in ABM, scenarios in CBMI) or tasks (e.g., dot-probe in ABM, ambiguous situations in CBMI) employed. We used the random effects model to account for the expected clinical and methodological heterogeneity and we observed no evidence to contradict our

assumption in terms of heterogeneity and inconsistency. Moreover, CBM tasks might differ in reliability and impact on targeted bias. We only investigated two study-level moderators, since, for a meaningful investigation of patient-level moderators, individual patient data would have been required (Debray et al., 2018).

The current network meta-analysis aggregates the largest number of CBM trials to date, and has the unique methodological advantage of simultaneously drawing from direct and indirect comparisons. CBMI emerged as promising and could warrant large-scale testing, ensuring blinding of participants and of outcome assessors and avoiding outcome reporting bias. Future trials would also need to clarify whether the intervention should be implemented as stand-alone or added to another, and whether it is cost-effective, given modest benefits observed so far.

This intervention could be construed as a schematic form of cognitive restructuring and it is unclear whether it deals with modifying an actual bias (i.e., implicitly) or, rather, prompts participants to explicitly use restructuring. Conversely, owing to the already large number of trials and small effects, it is doubtful whether further investment in the current dominant ABM paradigms as treatments is justified. This approach may only merit further experimentation in restricted settings with well-specified groups of subjects, at least until rigorous, pre-registered pre-clinical studies can demonstrate the reliability and benefits of alternative ABM paradigms.

3.3. Study 3: The effectiveness of a attentional bias assessment task in predicting symptoms of anxiety/depression

3.3.1. Introduction

Anxiety and depression disorders represent the most prevalent categories of mental health problems, with a minimum of 8.2% worldwide prevalence for anxiety disorders and 6.6% for depression, as of 2019 (Twenge & Joiner, 2020). Moreover, when confronted with disruptions to daily life, such as the recent COVID-19 pandemic, it has been shown that these numbers increase more than three-fold, 29.4% for anxiety and 24.9% for depression (Twenge & Joiner, 2020).

One paradigm proposes that deficits in neurocognitive functioning are associated with anxious/depressive symptomatology, with an abundance of evidence highlighting the role of impairments in attentional control (Eysenck et al., 2007; Pacheco-Unguetti et al., 2011; Rock et al., 2014). Effect sizes for comparisons between affected individuals and normal controls were 0.66 for anxiety and between 0.52 and 0.61 for depression (Pacheco-Unguetti et al., 2011; Rock et al., 2014).

Another paradigm related to attentional control brings forth evidence for attentional biases (ABM-eval) as having a causal and/or maintaining role in anxiety and depression disorders. Moreover, some aspects of attentional control (i.e., control of attentional inhibition and control of attentional selectivity) have been shown to be strongly related to the magnitude of attentional bias change (Basanovic et al., 2017). While the findings of early studies showed that attentional biases have a causal or maintaining role in anxiety disorders but not in depression (Dritschel, 1992; MacLeod et al., 1986), more recent studies have shown that attentional biases are also present in depression (Mogg et al., 1995; Peckham et al., 2010). More precisely, participants with anxious and/or depressive symptomatology tend to have an attentional bias toward negative stimuli when compared to normal controls (MacLeod et al., 1986; Mogg et al., 1995; Peckham et al., 2010). Two separate meta-analyses have revealed medium effect sizes for the comparison between anxious/depressed individuals and normal controls with regard to attention bias (Bar-Haim et al., 2007; Peckham et al., 2010). For anxious symptomatology, the effect size was $d = 0.45$ (Bar-Haim et al., 2007), while for depressive symptomatology the effect size was $d = 0.52$ (Peckham et al., 2010). On the other hand, there are studies showing that there are no significant associations between attentional biases and anxiety / depressive symptoms. For example, one study showed that none of the attention bias parameters predicted social anxiety symptoms in adolescents (Henricks et al., 2022). Another study revealed mixed findings with regard to the association between elevated trait anxiety symptoms and attentional biases, with some of the participants expressing an attention bias towards threat, some participants expressing an attention bias away from threat and some participants expressing an attention bias toward threat only for some categories of stimuli and away from threat for other categories of stimuli (Zvielli et al., 2014). In case of depressive symptoms, a recent study showed that there was no difference with regard to attention bias between individuals with major depressive disorder, dysphoric individuals and healthy individuals, regardless of the type of task or the type of stimuli that were used (Krings, 2020).

Taking all of the above into account our exploratory aim in the present study is to evaluate the effectiveness of a ABM evaluation procedure (ABM-eval) at discriminating between normal controls and individuals with anxious/depressive symptomatology.

3.3.2. Methods

Participants

Participants were recruited from 1) the general public (mainly students that received course credit in exchange for participation in the study). A total of 45 participants, aged between 19 and 52 (M = 27.73, SD = 8.44) took part in the study and were included in the analysis. Twenty-four percent were males (N = 11) and had a mean education of 16 years. Participants that were younger than 18 or had a history of neurological afflictions/substance dependence were excluded.

Based on the severity of anxious/depressive symptomatology, the sample was divided into healthy participants and participants with elevated symptomatology. The criteria that was used for this split was having above-cut-off scores on the Beck Depression Inventory (BDI-II, (Beck et al., 1996) and/or State-Trait Anxiety Inventory (Form Y, STAI-S, (Spielberger et al. 1983). More specifically, to be considered as having elevated symptoms of anxiety/depression, participants had to score equally or above 20 on the BDI-II and/or equally or above 34 on the STAI-S. These scores are considered to represent the threshold between moderate to severe symptoms and low to no symptoms according to the normative studies (Beck et al., 2012; Spielberger et al., 2007).

Attentional Bias Evaluation (ABM-eval)

The ABM-eval procedure followed the classical bias evaluation paradigm (MacLeod et al., 1986). The participants were instructed to look at the fixation cross that appears in the centre of the laptop display. After the fixation cross disappears, two paired faces, representing the same individual appear on the left and right of the screen, one face displaying a neutral expression and the other face displaying a disgust expression (i.e., neutral / threatening stimuli). The position of the neutral and threatening stimuli is randomised so that they appear with equal frequency on both sides of the screen. After 500 milliseconds, the faces disappear and a dot-probe appears in the place of one of the faces. The participants are instructed to indicate the location of the dot-probe as fast and as accurately as they can via keyboard input. There were 120 trials in total. As opposed to the classical intervention procedures, where the dot-probe replaces the neutral stimuli 80%-100% of the time, in the evaluation procedure the probe replaces the neutral and threatening stimuli with equal frequency (50%-50%). Attentional bias towards threat is considered to be present when response latencies are shorter for dot-probes that are located behind threatening stimuli as compared to neutral stimuli. The response latencies of the participants are recorded for each instance, that is for each repetition and an attentional bias score is computed by subtracting the average response time for neutral stimuli from the average response time for threatening stimuli.

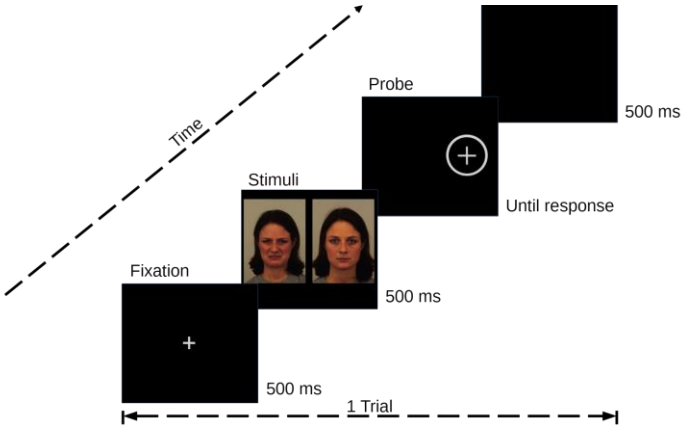

Figure 1. Example trial of a dot-probe task

Questionnaires

In addition to performing the attentional bias evaluation procedures, the participants also provided demographic characteristics (age, gender, medication taken at the time of the study). The symptoms of depression and anxiety were assessed with the BDI-II (Beck et al., 1996) or with STAI-S and STAI-T respectively (Spielberger et al., 1983).

Procedure

Participants were received either on the SkyRa platform at the the International Institute for the Advanced Studies of Psychotherapy and Applied Mental Health. All participants read and signed the informed consent form, followed by demographic data collection and they completed the BDI-II, STAI-S and STAI-T questionnaires. Following this, the ABM-eval procedure which was performed by the participants by performing the classic bias evaluation procedure, with all stimuli presented on a laptop screen.

Statistical analysis

In order to evaluate the effectiveness of ABM-eval in predicting anxious/depressive symptomatology we employed a logistical regression approach, the predictor being the attention bias score and the target being the status of the participants (anxious/depressive individuals or normal controls). We firstly computed the χ^2 metric in order to evaluate the overall statistical significance of the model. We employed the Nagelkerke R^2 indicator as an indicator of model-explained variance, while the confusion matrix, accuracy, area under the curve, sensitivity, specificity, the F1 score and precision were considered as indicators for model performance metrics. We also transformed the standardized coefficients into odds ratios (OR) and change in odds, for a more intuitive interpretation of the results, using the following equations: *OR = exp (StandardizedBeta)* and *%ChangeInOdds = (OR - 1) * 100*.

The attention bias score was computed by subtracting the mean reaction time of the participants toward disgust faces from the mean reaction time towards neutral faces. Thus, a positive bias index indicated that the participant reacted faster to

probes when they appeared behind neutral faces, while a negative bias index indicated a faster reaction to probes behind disgust faces:

BiasIndex = Mean (RTdisgust) - Mean (RTneutral)

3.3.3. Results

The participants had a mean age of 27.73 (SD = 8.44) years old. Seventy-six percent were females and 51% were employed. After participant segregation as a function of their clinical status, we had 19 participants with elevated levels of anxiety and/or depression and 26 healthy participants. There was a statistically significant difference between normal controls and individuals with anxious/depressive symptomatology, $t(43) = 3.19$, $p = 0.003$, Cohen's $d = 0.96$.

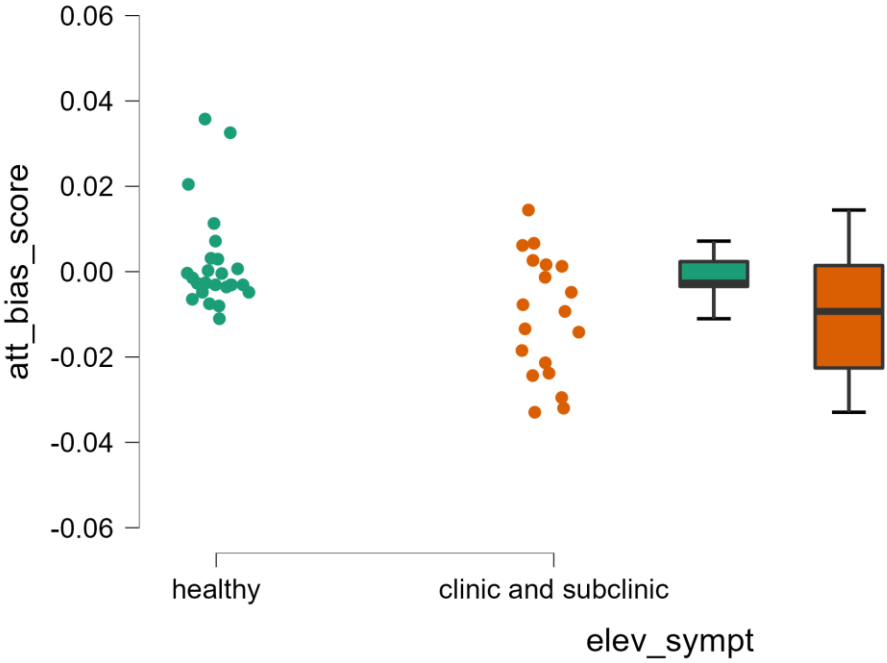

Figure 2. The distribution of attention bias scores as a function of clinical status

With regard to our hypothesis, namely to evaluate the effectiveness of ABM-eval at discriminating between normal controls and individuals with anxious/depressive symptomatology, the logistical regression model was significant, $\chi^2(43) = 10.29$, $p < 0.001$. Nagelkerke' R^2 indicated that 27.5% of the probability variance between normal controls and individuals with anxious/depressive symptomatology was explained by the ABM-eval attentional bias predictor which was statistically significant, *beta* = -1.28, p = 0.011. We obtained an OR = 0.27 and a %ChangeInOdds = -73%, meaning that each additional increase of attention bias to

threat is associated with an 73% decrease in the odds of a participant of not having symptoms of anxiety and/or depression.

Table 1. The regression coefficient for ABM-eval

Coefficients

					Wald Test		
	Estimate	**Standard Error**	**Standardized**[+]	**z**	**Wald Statistic**	**df**	**p**
(Intercept)	-0.722	0.378	-0.405	-1.908	3.642	1	0.056
att_bias_score	-91.752	35.956	-1.284	-2.552	6.512	1	0.011

Note. elev_sympt level 'clinic and subclinic' coded as class 1.

[+] Standardized estimates represent estimates where the continuous predictors are standardized (X-standardization).

The classification accuracy of the model was 75.55%, while the area under the curve was 0.70. We also observed high coefficients for specificity (0.92) and precision (0.83), while the coefficients for sensitivity (0.52) and the overall F1 indicator (0.64) were more modest

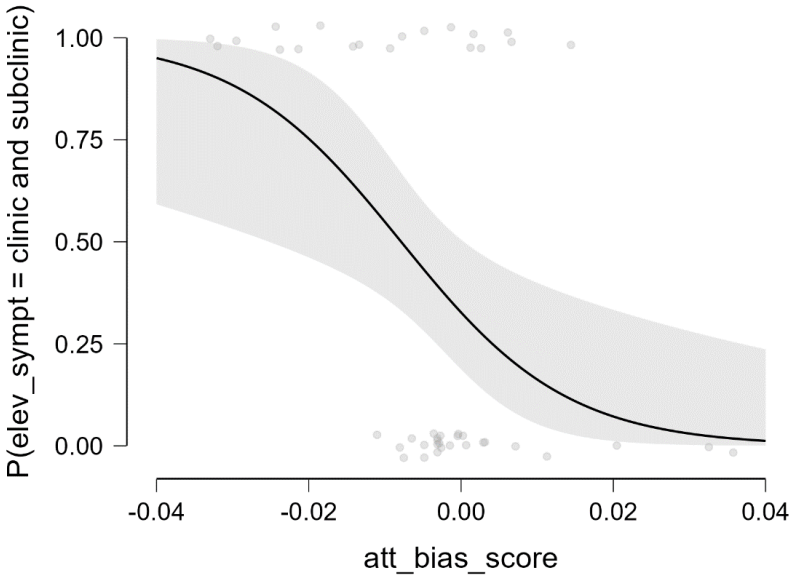

Figure 3. Conditional estimates plot

3.3.4. Discussion and conclusions

The exploratory aim of the present study was to evaluate the effectiveness of ABM-eval at discriminating between normal controls and individuals with anxious/depressive symptomatology. In this regard, we found that the attention bias-based method of evaluation of selective attention was effective at discriminating between normal controls and individuals with anxious/depressive symptomatology, ABM-eval having an strong explanatory power (27.5%). The classification coefficient of over 75% and an area under the curve of 0.70 indicated that the attention bias score represented a good fit for the model. Although the sensitivity score for the regression model was medium, the specificity and precision of the model were very high. These results confirm prior findings, in which the differences between healthy individuals and individuals with anxious/depressive symptomatology with regard to attention bias were Cohen's d = 0.45 for anxiety (Bar-Haim et al., 2007) and Cohen's d = 0.52 for depression (Peckham et al., 2010). The OR coefficient that we obtained is comparable to a Cohen's d of 0.72. This represents a larger difference between healthy individuals and individuals with anxious/depressive symptomatology than was observed in the aforementioned studies, one possible explanation for this result being the fact that our clinic and subclinic participant sample was an agglutination of individuals with anxious and depressive symptomatology.

The results that were obtained in the present study contradict the results that were observed in other studies (e.g., Henricks, 2022, Zvielli, 2014, Krings, 2020). However, given the fact that the previous meta-analytic research which aggregated a number of 172 (Bar-Haim et al., 2007) and 29 studies (Peckham et al., 2010) arrived at similar conclusions, the results of the present research can be considered to be consistent.

Another aspect that merits consideration in light of these results is the type of attention bias evaluation method. For example, one meta-analysis revealed that, while the effect size for the difference between healthy individuals and individuals with symptomatology was a statistically significant medium one in the case of those studies in which the dot-probe was used for attention bias evaluation, the effect size for the same difference was statistically non-significant for those studies in which the emotional Stroop task was usedfor attention bias evaluation. Given the fact that we used the classical dot-probe mode of attention bias evaluation, these results are even more consistent with the previous literature.

With regard to ABM-eval, its effectiveness in distinguishing between individuals with anxious/depressive symptomatology and normal controls, potentially driven by the specific stimuli that are used, opens new avenues of research. An evaluation method of this nature, when implemented in VR and benefitting from potential gamification, could represent an extremely viable, ecological and effective screening / evaluation method.

3.4. Study 4: The efficacy of Virtual Reality-Based Attention Bias Modification Training: A pilot randomized controlled trial

3.4.1. Introduction

Attentional bias evaluation and modification procedures (ABM) have been firstly developed and validated as early as 1986 (MacLeod et al., 1986). Since then, a lot of research has been conducted in order to evaluate the efficacy of ABM procedures in alleviating symptoms of anxiety (in particular with regard to social anxiety disorders), depression and addictions. The allure of these procedures consists in easy implementation and inexpensiveness, owing to the fact that they can be easily designed and have great compatibility with any computer system, while being easily administered, in an almost automatic fashion.

While there is a large body of research that investigates the efficacy of ABM procedures, both in modifying attentional biases and, through this mechanism alleviating anxious and/or depressive symptomatology, the results are often mixed. Meta-analyses conducted on this topic have shown that the effect size for bias modification is a moderate one and tends to get smaller after outlier removal (Cristea et al., 2015). When ABM is administered with the intention to reduce anxious and/or depressive symptomatology, the results are not very encouraging. For example, one meta-analysis revealed no statistically significant differences between ABM and control groups for either anxiety or depression (Cristea et al., 2015), while another more recent network meta-analysis revealed small effect sizes in favour of ABM for anxiety symptoms, but only in sensitivity analyses (Fodor et al., 2020).

With the adoption of virtual reality (VR) technology in augmenting (e.g., in vivo exposure) and in some cases replacing (e.g., attentional control evaluation) classical psychological therapeutic approaches, a new avenue of research is starting to be pursued for ABM interventions. VR implementations have many advantages over classical intervention delivery formats, advantages which were previously described in this thesis and not worth repeating. Presently, there are two studies that investigated the efficacy of VRABM (Ma et al., 2019; Urech et al., 2015). However, one of the studies was conducted as a proof-of-concept study (Urech et al., 2015), adopting a pre-post intervention design, without employing a control group. No statistically significant change was found from pre to post-intervention for attention bias or two out of three social anxiety measures. The other study (Ma et al., 2019) employed a complex randomized controlled trial (RCT) design, with two types of stimuli (2D vs 3D) and four groups (2D sham training, 2D ABM, 3D sham training and 3D ABM). Again, no "classical" intervention control group was employed, all groups experiencing VR immersion, the aim of the study being to discern the efficacy of different stimuli dimensionalities. No differences in attention bias were found either from pre to post-intervention or between experimental groups, while for anxiety measures only a time effect was present, anxiety decreasing over time across for all four groups.

Taking into account the small number of studies employing VRABM, the fact that no study compared a VRABM intervention to a "classical" ABM (PCABM) intervention and the fact that the small-to-moderate effect sizes in favour of ABM can be improved by VR adoption, it is clear that more research needs to be pursued in this direction. With this in mind, we conducted an pilot randomized controlled trial, in which we compared a VRABM active intervention to a PCABM active intervention, aiming at investigating the efficacy of the VR version as compared to the classical PC-delivered one in reducing attentional bias and improving potential state anxiety symptoms and fear of negative evaluation. We also aimed at evaluating the potential

adverse effects induced by VRABM, the sense of presence and the perceived usability of the VRABM intervention and, if the VRABM intervention induces more stress/mental workload as compared to PCABM. Because the superiority of PCABM over placebo and wait-list was previously established, we did not include these types of control groups in our study.

Owing to restrictions and limitations imposed by the COVID-19 pandemic (i.e., mobility restrictions, almost exclusive adoption of online methods of engagement and, as a direct consequence recruitment pool reduction), we were able to recruit only unselected participants through convenience sampling. This can be regarded as a limitation because usually, attentional biases towards threat are related to anxiety and participants with high anxiety levels exhibit the largest attentional bias towards threat. However, previous studies have shown that high threat attentional stimuli are capable of capturing attention in all participants, not only in those with above-threshold levels of anxious symptomatology (Mogg et al., 2000; Wilson & MacLeod, 2003). For the same reasons, the intervention consisted in only one session. However, previous studies have demonstrated that modifications in anxious symptomatology can and do occur in one session, both for ABM interventions (Ma et al., 2019; Sass et al., 2017) and for interpretation bias modifications (CBMI; (Beadel et al., 2016; Capron et al., 2017; Capron & Schmidt, 2016; Grisham et al., 2014; MacDonald et al., 2013; Mobini et al., 2014; Nowakowski et al., 2015; Steinman & Teachman, 2010; Vermeulen et al., 2019), as well as for attention bias (Amir et al., 2008; Buodo et al., 2018).

3.4.2. Method

Participants

Participants were recruited primarily via social media and through word-of-mouth. Taking into consideration the mobility restriction imposed due to the COVID-19 pandemic, especially at the beginning of the study, and in order to maintain a constant flow of participants, recruitment was restricted only to Cluj-Napoca. A total of 42 participants, aged between 15 and 52 (M = 28.16, SD = 6.49) took part in the study and were included in the analysis. Sixty-four percent were females (N = 27). Fifty percent of participants reported previous VR use, while 50% reported that they were using glasses or contact lenses. Participants that were younger than 18 or had a history of neurological afflictions/substance dependence or reported previous severe VR-induced motion sickness were excluded. We also excluded participants that failed to respond in the affirmative on each of the questions from the COVID-19 epidemiological questionnaire and/or had a bodily temperature greater than 37 degrees Celsius at screening.

Apparatus

VRABM was delivered via a HTC Vive head-mounted display (HMD) with a resolution of 2160×1200 (1080×1200 per eye), a field of view of 110 degrees and a refresh rate of 90 Hz. The HMD was paired with an ASUS Republic of Gamers laptop running on a Intel i7-8750H 2.20 Ghz processor, 24 Gb RAM and a GeForce GTX-1080 with 8 Gb VRAM video card. Participants interacted with stimuli by using the Vive controller that came with the HMD. PCABM was delivered on the aforementioned laptop, on screen, using a resolution of 1920x1080 and participants interacted with the stimuli using a mouse.

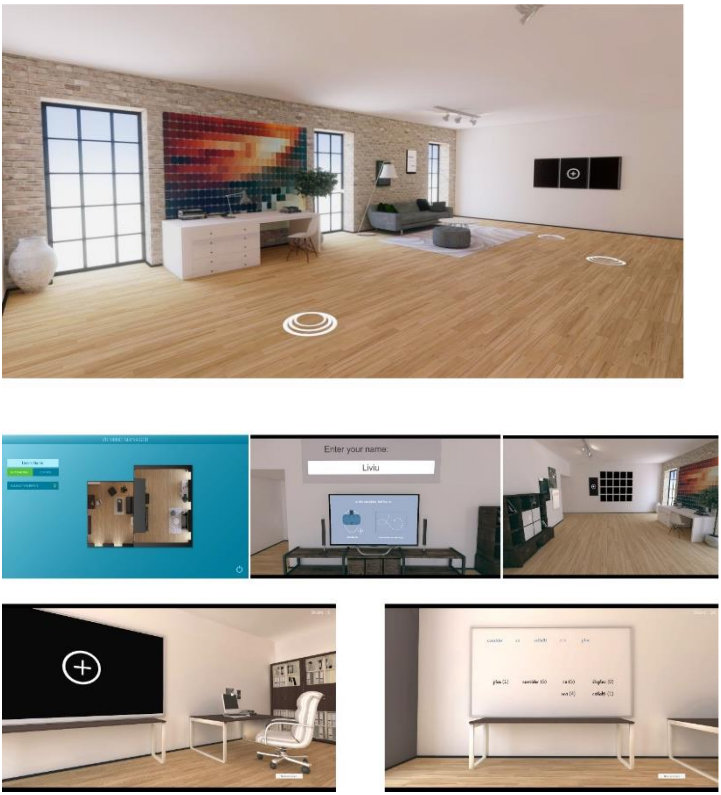

Figure 1. Screenshot of the VRABM environment

Attentional bias assessment and modification

The facial stimuli that we used for both VRABM and PCABM were selected from the Karolinska Directed Emotional Faces (Lundqvist et al., 1998). We selected 70 individuals, each showing a neutral and a negative-valence expression, of which 50% were female, for a total of 140 expressions. We chose disgust for the negative-valenced expressions because it is closely related to social anxiety (Amir et al., 2003; Phillips et al., 1998) and tends to elicit a higher number of complex emotions (i.e., humiliation, rejection and shame) than angry faces for example. The stimuli were identical across VRABM and PCABM and presented in the same environment. While in the PCABM condition the stimuli were presented on the laptop screen, in the VRABM condition the stimuli were presented on screen panels attached to a wall of the virtual room. In both conditions the participant made a dry, trial run, in order to get accommodated with the procedure.

We used the classical dot-probe task for both the pre-post measurement of attentional bias and the modification of attentional bias. There were a specific number of trials for each stage and each trial followed the following steps, also detailed in Figure 2: firstly, a fixation cross appeared in the center of the screen for 500 ms; secondly, after the fixation cross disappeared, two faces of the same individual were presented (one face depicting a neutral expression and the other depicting a disgust expression), arranged horizontally on the screen, for 500 ms; the position of the faces was counterbalanced, so that the disgust and neutral expressions appeared with equal frequency on the left or right side of the screen; thirdly, after the faces disappeared, a probe appeared in the location previously occupied by one of the faces, the probe position being also counterbalanced; fourthly, as previously instructed, participants reacted as fast as they could in indicating the position of the probe; finally, a 500 ms interval took place before a new trial would begin.

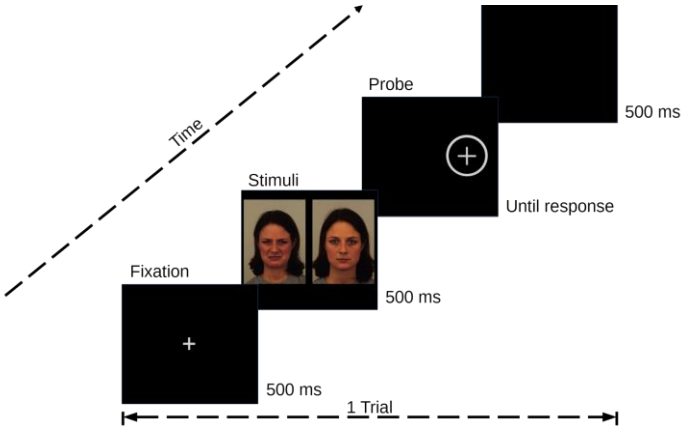

Figure 2. Example trial of a dot-probe task

For bias modification, the probe replaced in 80% of the cases the neutral expression as opposed to the disgust expression, while in 20% of the cases the pairs were neutral-neutral in order to mask the trial contingency. There were a total of 160 trials for both PCABM and VRABM conditions.

For bias measurement, a mock training design was used, in which the probe replaced words that had socially negative connotations / social-threat words (e.g., rejected, worthless, shameful, marginalized, humiliated, criticised, embarrassed) or words that had socially neutral connotations (e.g., amazed, temporary, ongoing, country, original, governmental), with equal frequency (50%-50%). There were 16 social-threat words and 16 neutral words that were matched for length. We chose to use words instead of faces for measuring attention bias in order to firstly avoid task habituation and secondly, to avoid any carryover effects from bias modification to bias measurement, especially at the post-intervention assessment.

In both VRABM and PCABM interventions, the same structure was followed: firstly, the pre-intervention bias score assessment was made, followed by the intervention proper. Lastly, the post-intervention bias score was assessed.

Bias assessments were made by subtracting the mean reaction time of the participants toward disgust faces from the mean reaction time towards neutral faces. Thus, a positive bias index indicated that the participant reacted faster to probes when they appeared behind neutral faces, while a negative bias index indicated a faster reaction to probes behind disgust faces:

BiasIndex = Mean (RTdisgust) - Mean (RTneutral)

Questionnaires

The measures that were specific only for VRABM were 1) simulator sickness (i.e., adverse effects) and was assessed pre and post exposure to VRABM using the Simulator Sickness Questionnaire - SSQ (Kennedy et al., 1993), and 2) system usability and the level of sense of presence in VR that were assessed at post-intervention with the System Usability Scale - SUS (Brooke, 1996) and the Presence Questionnaire – PQ (Witmer & Singer, 1998), respectively.

The measures that were common to both VRABM and PCABM were 1) state anxiety symptoms that were assessed both at pre and post intervention with the State-Trait Anxiety Inventory - STAI-S (Spielberger et al. 1983; cut-off: 40), 2) the fear of negative evaluation that was assessed both at pre and post intervention with the Brief Fear of Negative Evaluation scale (Leary, 1983; range: 12-60), and 3) the stress and mental workload was assessed by using the NASA Task Load Index tool - NASA-TLX (Hart & Staveland, 1988).

Procedure

Participants were received on the SkyRa platform at the International Institute for the Advanced Studies of Psychotherapy and Applied Mental Health. A thorough disinfection procedure was designed and implemented in order to prevent SARS-Cov2 infections and, to this end, before receiving a participant and after the participant had left a thorough disinfection of all surfaces and apparatus was conducted with Hexasept, a coronavirus virucide. Moreover, both the participants and the experimenter wore facial masks. The participants' temperature was taken and they completed the epidemiological triage questionnaire, both of which were used as primary inclusion/exclusion criteria. Participants were then assigned, based on a previously generated random sequence (generated at www.random.org) to either VRABM or PCABM. The VRABM participants completed the informed consent, the demographic data questionnaire, the pre-intervention SSQ, STAI-S and BFNE questionnaires and underwent pre-intervention bias assessment, bias modification and post-intervention bias assessment. Following this, the participants in the VRABM group completed the post-intervention SSQ, STAI-S, BFNE questionnaires and the NASA-TLX, PQ and SUS questionnaires. The participants in the PCABM group followed the same sequence, with the exception that they did not have to complete the pre/post intervention SSQ questionnaire and the PQ and SUS questionnaires.

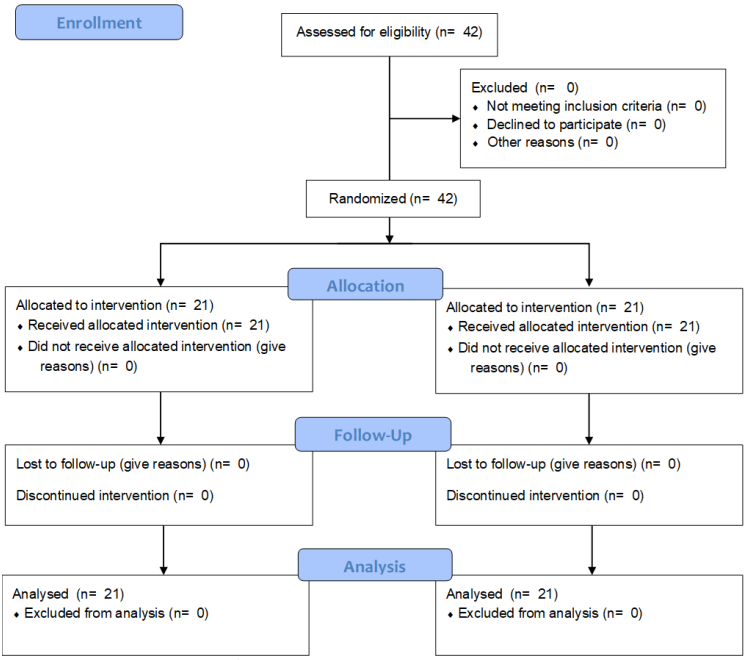

Figure 3. The CONSORT flow chart illustrating the randomisation process

Statistical analysis

Firstly, we ran a descriptive analysis of the data in order to evaluate if there were missing observations or extreme outliers and to check the distribution shapes for all variables. In order to estimate the results, we employed a 3 (intervention group: VRABM, PCABM) x 2 (assessments: pre-intervention, post-intervention) mixed-design ANCOVA, using baseline scores as covariates. This approach has been chosen because using pre-intervention values as a covariate provides more statistical power and more precise confidence intervals with regard to intervention effects than a mixed-design ANOVA (Rausch et al., 2003; "The Oxford Handbook of Research Strategies for Clinical Psychology," 2013). The assumption of homogeneity of variances was tested by employing the Levene test and possibly significant group interactions were followed with Tukey-adjusted pairwise comparisons of estimated marginal means. For the estimation of the effect size in the case of main group effects, $\eta_2 p$ was computed, while for significant pairwise comparisons, Cohen's d was employed. For stress/mental workload, that was measured only at post-intervention in both groups, we performed an independent-samples t-test. For variables that were measured only at post-intervention in the VRABM group (PQ, SUS), we interpreted the results based on range (i.e., median, minimum, maximum), mean and standard deviations. We also computed the Reliable Change Index (RCI) for each participant in both groups to further explore possible changes ain attentional bias and in order to see more clearly how many

individual participants deteriorated, improved or did not suffer any changes as a result of the interventions.

3.4.3. Results

The characteristics of the included participants are detailed in Table 1. The mean age was 28.16 years old and 64% of the participants were female. There were no drop-outs from the any of the intervention groups and there were no differences at baseline with regard to any of the investigated variables.

Table 1. Characteristics of the participants at baseline

		VRABM (N = 21)	PCABM (N = 21)	
Female	N (%)	13 (61.90)	14 (66.66)	$\chi^2(1) = 0.104$, $p = 0.747$
Having vision correction	N (%)	10 (47.61)	11 (52.38)	$\chi^2(1) = 0.09$, $p = 0.758$
Previous VR exposure	N (%)	12 (57.14)	9 (42.85)	$\chi^2(1) = 0.85$, $p = 0.355$
Age	M (SD)	27.85 (4.79)	28.47 (7.95)	$t(40) = 0.30$, $p = 0.762$
STAI-S pre-intervention	M (SD)	29.04 (6.60)	31.61 (7.66)	$t(40) = 1.16$, $p = 0.251$
BFNE pre-intervention	M (SD)	31.57 (8.50)	32.04 (9.56)	$t(40) = 0.17$, $p = 0.866$

With regard to the first objective, namely to investigate the efficacy of VRABM as compared to PCABM in reducing attentional bias and improving potential state anxiety symptoms and fear of negative evaluation, the results were mixed. The Levene's test was statistically non-significant for all three outcomes, so the homogeneity of variances assumption was met. For attention bias, the ANCOVA analysis revealed no statistically significant differences between groups: $F(1, 39) = 0.43$, $p = 0.514$. However, the RCI analysis revealed that there were more improvements and less deteriorations with regard to attentional bias in the VRABM group than in the PCABM group (VRABM: 12 improved, 5 deteriorated and 4 no change; PCABM: 7 improved, 9 deteriorated and 5 no change).

Also, no statistically significant differences were found with regard to the fear of negative evaluation, $F(1, 39) = 0.04$, $p = 0.839$. With regard to state anxiety symptoms, we identified a statistically significant difference between the VRABM and PCABM groups: $F(1, 39) = 74.20$, $p = 0.016$, $\eta^2p = 0.07$. Post-hoc analyses revealed that the effect size was medium, Cohen's $d = 0.55$ and the estimated marginal means were 27.24 (95% CI: 25.72 to 28.76) for VRABM and 29.94 (95% CI: 28.42 to 31.47) for PCABM.

In terms of potential adverse effects induced by the VRABM intervention, the participants in the VRABM group reported little to no symptoms post-VR exposure. The participants in the VRABM condition had a M = 2.33, SD = 1.82 at pre-intervention, while at post-intervention they had a M = 2.81, SD = 2.80 (possible score range: 0 – 48, median value: 24). A paired sample t-test revealed no statistically significant difference from pre- to post-VR-intervention, $t(20) = -0.76$, $p = 0.454$.

With regard to stress/mental workload the assumptions of normality and of equality of variances were met (VRABM: Shapiro-Wilk W = 0.972, p = 0.776; PCABM: Shapiro-Wilk W = 0.973, p = 0.795; Levene's F = 0.35, p = 0.555). The participants in the PCABM group reported statistically significant lower stress/mental workload (M = 159.52, SD = 61.05) than the participants in the VRABM group (M = 205.71, SD = 70.98), $t(40) = -2.26$, $p = 0.029$, Cohen's d = -0.69 However, taking into account the possible score range of 0 – 600, with a median value of 300, stress/mental workload was low in both groups.

With regard to presence in VR, the participants in the VRABM group reported medium to high levels of presence (M = 168.76, SD = 16.12; possible score range: 32 – 224, median value: 128). In terms of system usability, the participants in the VRABM group rated the VR system as having an above average-to-excellent usability, the mean score being M = 86.55, SD = 9.94 (i.e., a score above 68 out of a range from 10 to 100).

3.4.4. Discussion and conclusions

The aims of the present study were 1) to evaluate the efficacy of the VRABM intervention as compared to the PCABM intervention in reducing attentional bias and improving potential state anxiety symptoms and fear of negative evaluation, 2) to evaluate if the VRABM intervention induces more stress/mental workload as compared to the PCABM intervention and 3) to evaluate the potential adverse effects induced by the VRABM intervention, the participants' sense of presence in VR and the perceived usability of the VRABM intervention. With regard to the first aim, we found that there was no statistically significant difference between VRABM and PCABM in reducing attentional bias. This result could, in theory, be attributed to two factors: 1) the participants were unselected, both in terms of state anxiety and of fear of negative evaluation, meaning that although the scores, as a mean, extended in the subclinical range for both instruments, they never reached the cut-offs for clinical symptomatology and 2) we compared two active interventions, both of which aimed at reducing the aforementioned symptoms and, although equivalence between interventions cannot be claimed based on a statistically non-significant result, this could still represent a factor. Both arguments are supported, when taking into consideration that, in previous studies, attentional bias did not change even when the ABM VR intervention was compared to an active or placebo classical intervention (Ma et al., 2019), when change in bias was evaluated only from pre- to post-intervention (Urech et al., 2015), when subclinical samples were employed or when single session interventions were employed (Ma et al., 2019; Urech et al., 2015). However, with regard to attentional bias, the RCI analysis revealed that 12 participants improved and 4 did not change in the VRABM condition, while only 7 improved and 5 did not change in the PCABM condition. Moreover, only 5 participants deteriorated with regard to attentional bias in the VRABM condition, while 9 participants deteriorated in the PCABM condition. This represents an encouraging result, albeit anecdotal, with regard to the superiority of the VRABM intervention over the PCABM intervention.

The same argument as in the case of attention bias change can be applied with regard to fear of negative evaluation symptoms, a component of social anxiety, for which we did not found any statistically significant differences. In the two previous VR studies, changes in social anxious symptomatology were not detected when measured with the Social Phobia Scale or with the Social Interaction Anxiety Scale (Urech et al., 2015) or with the Liebowitz Social Anxiety Scale (Ma et al., 2019). Urech et al., (2015) found a small effect from pre- to post-intervention on the Liebowitz Social Anxiety Scale, but it can be arguably justified as statistical artefact, since on the other two social anxiety scales that were employed, no such effect was identified. Moreover, fear of negative evaluation in special and social anxiety in general represent constructs that may not be so prone to modification in only one-session interventions, even when employing more ecological methodologies, such as VR.

This is theoretically not the case for state anxiety, which is more prone to modifications in single-session interventions, as evidenced by the statistically significant effect that was obtained between VRABM and PCABM ($d = 0.55$). Although it is conceivable that this result can represent a statistical artefact especially when considering a one-session intervention and the small sample size, this result was observed in previous studies ($d = 0.45$ in Amir et al., 2008; $d = 0.48$ in Dennis & O'Toole, 2014), which leads credence to the superiority of VRABM over PCABM in reducing state anxiety symptoms.

With regard to stress/mental workload, the participants in the VRABM group reported having a mode challenging / demanding experience than the participants in the PCABM group, although the scores remained well below the median of the scale. This result was to be expected, when taking into account the fact that the virtual environment, through properties such as enhanced spatiality, sense of presence, near isolation to normal external stimuli and the usage of special joysticks, makes it more challenging for the user to perform the ABM task, especially when said user has not been acquainted with any form of VR technology previously. This is not the case for the participants in the PCABM condition, as they had only to sit in front of a laptop and perform the task by using the mouse, a situation that would be familiar to almost any person.

Finally, with regard to sense of presence in VR and perceived system usability, the participants in the VRABM group reported high levels for both measures. Given the fact that the sense of presence in VR represents a factor that is directly involved in VR treatments' efficacy (Wallach et al., 2012), this result is very encouraging and leads credence to VRABM as an potentially effective intervention in longer time-frames (i.e., multiple VRABM sessions). Similarly, the above average-to-excellent usability, supports the fact that the VRABM system is easy to learn, with minimal instructions and the task performance is not impeded by any properties of the hardware platform or, more importantly by the design of the VRABM intervention.

CHAPTER IV. GENERAL CONCLUSIONS AND IMPLICATIONS

4.1. General Conclusions

We aimed to address a number of methodological objectives related to attentional bias evaluation and virtual reality mediated attentional bias modification interventions in this thesis. More specifically, we wanted to elucidate if the introduction of new, virtual reality-based procedures for evaluating attentional biases and modifying attentional biases were at least as efficient as classical computerised techniques. In order to accomplish these objectives, a number of intermediary steps had to be performed, reflected in our original research articles.

First, a systematic review and meta-analysis was conducted, of virtual reality interventions for anxiety and comorbid depression outcomes, as well as treatment attrition. We included randomized controlled trials comparing VR interventions, alone or in combination, to control conditions or other active psychological interventions. The main findings were that VR-based therapies were more effective than passive controls at post-test for anxiety and comorbid depression, but not for treatment attrition. We also revealed that the effect sizes were higher when participants were recruited from a clinical setting, or when the diagnostics were specific phobia, panic disorder, flight anxiety or social phobia. Moreover, in the contrast with other active interventions, the effect size was higher for virtual-reality-based exposure than for virtual reality-based cognitive-behavioural therapy. Also, the number of elements of interaction with the virtual environment was positively associated with anxiety outcomes, a result that lends additional credence to the importance of immersion and presence concepts. There were no significant differences between virtual reality-based interventions and other active interventions.

Second, we conducted a systematic review and network meta-analysis, in which we evaluated the relative effectiveness of CBM procedures (i.e., ABM, CBMI, AAT), as simultaneously compared among each other and with various control groups, for anxious and depressive symptomatology, as well as for comorbid anxious and depressive symptomatology. We included randomized controlled trials comparing a cognitive bias modification intervention to a control condition for anxious or depressive symptomatology, as measured on validated clinical scales, in adults whose primary complaint consisted of symptoms of anxiety or depression. For anxious symptomatology, only the contrast between interpretation bias modification and waitlist or the contrast between interpretation bias modification and placebo were significant. For depression outcomes, again, only the contrast between interpretation bias modification and waitlist was significant, together with the contrast between the combined treatments (attention plus interpretation bias modification) and waitlist. For comorbid depression outcomes in anxiety trials, the contrast between interpretation bias modification and waitlist or the contrast between interpretation bias modification and placebo were significant. For comorbid anxiety symptoms in depression trials, there were no significant results. The attention bias modification interventions were superior to placebo and waitlist only in sensitivity analyses, in which trials concerned with posttraumatic stress disorder symptomatology were excluded. The modest results for attention bias modification interventions, together with the mixed results obtained in previous meta-analyses, lead credence to the theory that newer, more interactive (i.e., virtual reality) implementations might be needed in order to improve on these outcomes.

Third, we conducted a cross sectional experimental study in which we evaluated the effectiveness of a attentional bias evaluation procedure, at discriminating

between healthy individuals and individuals with anxious/depressive symptomatology. The attention bias evaluation procedure was not only efficient at discriminating between anxious / depressive participants from healthy ones, but also accurate, thus confirming previous findings from the literature, and bringing additional validity to the dot-probe methods' efficacy at evaluating attentional bias.

Fourth, we conducted a randomized controlled trial in which we evaluated the efficacy of a newly developed, virtual reality-based attention bias modification procedure, as compared to the classical computerized procedure at reducing attentional bias, anxious symptomatology. We also investigated other aspects, relevant for the virtual reality environment, such as usability, adverse effects, stress/perceived mental workload and the level of presence in virtual reality. While the results for attentional bias score and fear of negative evaluation were not significant, we observed a significant reduction in state anxiety in favour of the virtual reality group. Also, as a tentative result, more participants improved and less deteriorated in the virtual reality group than in the computerized group. The participants in the virtual reality group reported high levels of presence and excellent virtual system usability. Moreover, stress/mental workload was low in both experimental groups, although the participants in the virtual reality group reported significantly more stress/mental workload. There were virtually no adverse effects as a result of the virtual reality intervention.

Summarising, through the present work we found out that 1) virtual reality interventions are superior to controls for both anxious and comorbid depressive symptomatology, 2) attention bias modification procedures are superior to placebo and waitlist for both anxious and depressive symptomatology, in certain conditions, 3) the dot-probe-based attentional bias evaluation discriminated accurately between anxious / depressive participants and healthy ones and 4) that virtual reality-based attention bias modification represents a feasible intervention with tentative results that merit further investigation. These findings compel us to consider a number of methodological and clinical consequences, which are described further below.

4.2. Implications of the present thesis

4.2.1. Methodological implications

From a methodological point of view, the present thesis brings some contributions and fills some gaps in the literature with regard to virtual reality-based evaluation methods and interventions. More specifically, through the first study we updated the methodology regarding virtual reality by taking into account the latest studies in the literature and addressing gaps in previous meta-analytical approaches. More specifically, some of the unique contributions of this study are that we updated the list of included studies to reflect the latest research in the virtual reality, we investigated the effects of virtual reality-based interventions on comorbid depression, we investigated the effects of virtual reality interventions on treatment attrition and investigated the effects of previously untested potential moderators.

The second study represents in our opinion a significant methodological advance. It is the first network meta-analysis in the field of cognitive bias modification literature. Through this new methodological approach, we were able to investigate the efficacy all types of cognitive bias modification procedures simultaneously and derive indirect contrast between interventions that have never been directly compared before, either in a randomized controlled trial or in a meta-analytical approach. Moreover, we have taken into consideration not only anxious and depressive symptomatology, but also comorbid anxious and depressive symptomatologies.

The third study brings forward some methodological advances, namely the fact that we tested the efficacy of a attentional bias evaluation procedure at discriminating between anxious / depressive symptomatologies and healthy controls. The novelty here is represented by the fact that we demonstrated and confirmed the viability of the dot-probe task at discriminating between healthy and non-healthy individuals and, classifying the cases with a relatively high degree of accuracy.

Finally, the fourth study's main methodological contribution is represented by the fact that the virtual reality-based cognitive bias modification procedures that we used represent a novel and significant addition to the two existing (Ma et al., 2019; Urech et al., 2015) virtual reality attention bias modification interventions. The virtual environment and the evaluation / modification procedures were envisioned and designed by the author of this thesis and Silviu Matu, Ph.D., under the supervision of Professor Daniel David and developed by the E.ON Reality software company, from the ground up. Although not evaluated in the present thesis, the virtual reality-based cognitive bias modification software includes not only the attention bias evaluation and modification procedure, but also a memory bias modification and a interpretation bias modification procedure.

4.2.2. Clinical implications

In addition to the methodological implications, a series of clinical implications can be derived from the present thesis. Mainly, through finding that virtual reality-based interventions are effective in reducing symptoms of anxiety and comorbid depression as compared to control in Study 1, an avenue is opened not only to patients but also to practitioners in using this evidence-based approach in managing anxious / depressive symptomatology, especially when knowing that this type of technology has been proved to be cost effective (Wood et al., 2009; Freeman et al., 2017). Moreover, both patients and therapists can make informed treatment decisions, knowing that the best results are obtained for specific diagnoses (i.e., specific phobia, panic disorder, flight anxiety, social phobia), especially when virtual reality exposure is used. In addition, companies that develop these types of technologies with the specific aim of therapy use can be informed by the fact that number of elements of interaction with the virtual environment was positively associated with better anxiety outcomes. The main clinical implication stemming from Study 2, especially relevant for therapists employing cognitive bias modification procedures either as standalone or as an adjuvant intervention, is that it is recommendable to employ interpretation bias modification procedures at the least, or a combination of attention bias modification and interpretation bias modification procedures at the most, when managing anxious / depressive symptomatology. Study 3 also has some clinical implications, the main ones being that the dot-probe-based attentional bias assessment represents a procedure that has superior discriminatory power between healthy and clinical populations, thus having the potential to be used as an alternative screening tool, being also mature enough as to be a good candidate for virtual reality implementation. The main clinical contribution of Study 4 is that it shown that a virtual reality-based attention bias modification procedure, at the minimum, has the potential of reducing state anxiety and can be used in this sense as an adjuvant to classical evidence-based therapies.

4.3. Limitations and Further Avenues of Research

As is the case with any research approach, the present thesis has a series of limitations worthy of being mentioned. First, Study 1 revealed high degrees of heterogeneity and large confidence intervals, which add limitations to the solidity of the results. Also, many of the subgroup analyses were underpowered and few moderators were reported in the primary studies. Second, in Study 2 most trials offered no information in relation to study quality, most trials being rated as unclear for most study quality domains. Another limitation of Study 2 is represented by clinical heterogeneity (placebo treatments were considered to be interchangeable across interventions because the principle behind them is the same). Taken together, as in the case of Study 1, these limitations suggest using caution when drawing clinical conclusions. Study 3 also has a series of limitations, the most important one being that we could not separate anxious and depressive symptomatology to run separate comparative analyses, owing to the fact that such a segregation would have lead to underpowered logistical regression models. Moreover, for the same reason, we considered the clinical and subclinical participants as being part of the same group. With regard to Study 4, the main limitation is that we had a small sample size and the comparative analysis between the virtual reality group and the classical intervention group are underpowered. This limitation was mainly imposed by the apparition of the SARS-Cov-2 virus, the study being conducted in the very first year of the epidemic. Another potential limitation of Study 4 is that we did not use clinical samples which could have been more appropriate for the aim of the intervention that were delivered. Taken together, these limitations are indicative of further improvements that can be made in future research, mainly related to recruitment procedures, increased sample sizes and participants clinical status which, if pursued, could make a great difference in the generalizability of the research.

REFERENCES

Abraha, I., Cherubini, A., Cozzolino, F., Florio, R. D., Luchetta, M. L., Rimland, J. M., Folletti, I., Marchesi, M., Germani, A., Orso, M., Eusebi, P., & Montedori, A. (2015). Deviation from intention to treat analysis in randomised trials and treatment effect estimates: Meta-epidemiological study. BMJ, 350, h2445. https://doi.org/10.1136/bmj.h2445

Amad, A., Jardri, R., Rousseau, C., Larochelle, Y., Ioannidis, J. P. A., & Naudet, F. (2019). Excess Significance Bias in Repetitive Transcranial Magnetic Stimulation Literature for Neuropsychiatric Disorders. Psychotherapy and Psychosomatics, 88(6), 363–370. https://doi.org/10.1159/000502805

Amir, N., Elias, J., Klumpp, H., & Przeworski, A. (2003). Attentional bias to threat in social phobia: Facilitated processing of threat or difficulty disengaging attention from threat? Behaviour Research and Therapy, 41(11), 1325–1335. https://doi.org/10.1016/s0005-7967(03)00039-1

Amir, N., Weber, G., Beard, C., Bomyea, J., & Taylor, C. T. (2008). The Effect of a Single-Session Attention Modification Program on Response to a Public-Speaking Challenge in Socially Anxious Individuals. Journal of Abnormal Psychology, 117(4), 860–868. https://doi.org/10.1037/a0013445

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. https://doi.org/10.1037/a0033559

Andrews, G., Basu, A., Cuijpers, P., Craske, M. G., McEvoy, P., English, C. L., & Newby, J. M. (2018). Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: An updated meta-analysis. Journal of Anxiety Disorders, 55, 70–78. https://doi.org/10.1016/j.janxdis.2018.01.001

Armijo-Olivo, S., Fuentes, J., Ospina, M., Saltaji, H., & Hartling, L. (2013). Inconsistency in the items included in tools used in general health research and physical therapy to evaluate the methodological quality of randomized controlled trials: A descriptive analysis. BMC Medical Research Methodology, 13(1), 116. https://doi.org/10.1186/1471-2288-13-116

Azriel, O., & Bar-Haim, Y. (2020). Attention bias. In Clinical handbook of fear and anxiety: Maintenance processes and treatment mechanisms (pp. 203–218). American Psychological Association. https://doi.org/10.1037/0000150-012

Baddeley, A. (1998). The central executive: A concept and some misconceptions. Journal of the International Neuropsychological Society, 4(5), 523–526. https://doi.org/10.1017/S135561779800513X

Badura-Brack, A. S., Naim, R., Ryan, T. J., Levy, O., Abend, R., Khanna, M. M., McDermott, T. J., Pine, D. S., & Bar-Haim, Y. (2015). Effect of Attention Training on Attention Bias Variability and PTSD Symptoms: Randomized Controlled Trials in Israeli and U.S. Combat Veterans. The American Journal of Psychiatry, 172(12), 1233–1241. https://doi.org/10.1176/appi.ajp.2015.14121578

Bar-Haim, Y. (2010). Research review: Attention bias modification (ABM): a novel treatment for anxiety disorders. Journal of Child Psychology and Psychiatry, and Allied Disciplines, 51(8), 859–870. https://doi.org/10.1111/j.1469-7610.2010.02251.x

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. Psychological Bulletin, 133(1), 1–24. https://doi.org/10.1037/0033-2909.133.1.1

Basanovic, J., Notebaert, L., Grafton, B., Hirsch, C. R., & Clarke, P. J. F. (2017). Attentional control predicts change in bias in response to attentional bias modification. Behaviour Research and Therapy, 99, 47–56. https://doi.org/10.1016/j.brat.2017.09.002

Beadel, J. R., Ritchey, F. C., & Teachman, B. A. (2016). Role of Fear Domain Match and Baseline Bias in Interpretation Training for Contamination Fear. Journal of Experimental Psychopathology, 7(1), 49–71. https://doi.org/10.5127/jep.045414

Beck, A. T., David, D., & Dobrean, A. (2012). Inventarul de depresie Beck-Editia a doua (BDI-II). Romanian Testing Services.

Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. (1996). Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. Journal of Personality Assessment, 67(3), 588–597. https://doi.org/10.1207/s15327752jpa6703_13

Bernstein, A., & Zvielli, A. (2014). Attention Feedback Awareness and Control Training (A-FACT): Experimental test of a novel intervention paradigm targeting attentional bias. Behaviour Research and Therapy, 55, 18–26. https://doi.org/10.1016/j.brat.2014.01.003

Blackwell, S. E., Woud, M. L., & MacLeod, C. (2017). A Question of Control? Examining the Role of Control Conditions in Experimental Psychopathology using the Example of Cognitive Bias Modification Research. The Spanish Journal of Psychology, 20. https://doi.org/10.1017/sjp.2017.41

Boettcher, J., Leek, L., Matson, L., Holmes, E. A., Browning, M., MacLeod, C., Andersson, G., & Carlbring, P. (2013). Internet-based attention bias modification for social anxiety: A randomised controlled comparison of training towards negative and training towards positive cues. PloS One, 8(9), e71760. https://doi.org/10.1371/journal.pone.0071760

Booth, R., & Rachman, S. (1992). The reduction of claustrophobia—I. Behaviour Research and Therapy, 30(3), 207–221. https://doi.org/10.1016/0005-7967(92)90067-Q

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). Introduction to Meta-Analysis (1st ed.). Wiley.

Botella, C., Baños, R. M., Perpiñá, C., Villa, H., Alcañiz, M., & Rey, A. (1998). Virtual reality treatment of claustrophobia: A case report. Behaviour Research and Therapy, 36(2), 239–246. https://doi.org/10.1016/S0005-7967(97)10006-7

Botella, C., Serrano, B., Baños, R. M., & Garcia-Palacios, A. (2015). Virtual reality exposure-based therapy for the treatment of post-traumatic stress disorder: A review of its efficacy, the adequacy of the treatment protocol, and its acceptability. Neuropsychiatric Disease and Treatment, 11, 2533–2545. https://doi.org/10.2147/NDT.S89542

Bouchard, S., Dumoulin, S., Robillard, G., Guitard, T., Klinger, É., Forget, H., Loranger, C., & Roucaut, F. X. (2017). Virtual reality compared with in vivo exposure in the treatment of social anxiety disorder: A three-arm randomised controlled trial. The British Journal of Psychiatry, 210(4), 276–283. https://doi.org/10.1192/bjp.bp.116.184234

Brooke, john. (1996). SUS: A "Quick and Dirty" Usability Scale. In Usability Evaluation In Industry. CRC Press.

Brown, C., Schulberg, H. C., Madonia, M. J., Shear, M. K., & Houck, P. R. (1996). Treatment outcomes for primary care patients with major depression and lifetime anxiety disorders. The American Journal of Psychiatry, 153(10), 1293–1300. https://doi.org/10.1176/ajp.153.10.1293

Buodo, G., Patron, E., Messerotti Benvenuti, S., & Palomba, D. (2018). Single-Session Attention Bias Modification Training in Victims of Work-Related Accidents. Frontiers in Psychology, 9, 1619. https://doi.org/10.3389/fpsyg.2018.01619

Campbell, M., & Kertz, S. (2019). Brief report: Attentional Control Moderates the Relationship Between Attentional Bias and Anxiety in Children. Child & Family Behavior Therapy, 41(1), 1–15. https://doi.org/10.1080/07317107.2019.1571768

Capron, D. W., & Schmidt, N. B. (2016). Development and randomized trial evaluation of a novel computer-delivered anxiety sensitivity intervention. Behaviour Research and Therapy, 81, 47–55. https://doi.org/10.1016/j.brat.2016.04.001

Capron, D. W., Norr, A. M., Allan, N. P., & Schmidt, N. B. (2017). Combined "top-down" and "bottom-up" intervention for anxiety sensitivity: Pilot randomized trial testing the additive effect of interpretation bias

modification. Journal of Psychiatric Research, 85, 75–82. https://doi.org/10.1016/j.jpsychires.2016.11.003
Chaimani, A., Higgins, J. P. T., Mavridis, D., Spyridonos, P., & Salanti, G. (2013). Graphical Tools for Network Meta-Analysis in STATA. PLOS ONE, 8(10), e76654. https://doi.org/10.1371/journal.pone.0076654
Cheng, J., Pullenayegum, E., Marshall, J. K., Iorio, A., & Thabane, L. (2016). Impact of including or excluding both-armed zero-event studies on using standard meta-analysis methods for rare event outcome: A simulation study. BMJ Open, 6(8), e010983. https://doi.org/10.1136/bmjopen-2015-010983
Christensen, E. (2007). Methodology of superiority vs. Equivalence trials and non-inferiority trials. Journal of Hepatology, 46(5), 947–954. https://doi.org/10.1016/j.jhep.2007.02.015
Comeau, C. (1961). Headsight Television System Provides Remote Surveillance. Electronics, 86–90.
Comer, J. S., & Kendall, P. C. (Eds.). (2013). The Oxford handbook of research strategies for clinical psychology (pp. xvi, 459). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199793549.001.0001
Conn, C., Lanier, J., Minsky, M., Fisher, S., & Druin, A. (1989). Virtual environments and interactivity: Windows to the future. ACM SIGGRAPH Computer Graphics, 23(5), 7–18. https://doi.org/10.1145/77277.77278
Cristea, I. A. (2019). The waiting list is an inadequate benchmark for estimating the effectiveness of psychotherapy for depression. Epidemiology and Psychiatric Sciences, 28(3), 278–279. https://doi.org/10.1017/S2045796018000665
Cristea, I. A., Kok, R. N., & Cuijpers, P. (2015). Efficacy of cognitive bias modification interventions in anxiety and depression: Meta-analysis. The British Journal of Psychiatry: The Journal of Mental Science, 206(1), 7–16. https://doi.org/10.1192/bjp.bp.114.146761
Cristea, I. A., Kok, R. N., & Cuijpers, P. (2016). The Effectiveness of Cognitive Bias Modification Interventions for Substance Addictions: A Meta-Analysis. PloS One, 11(9), e0162226. https://doi.org/10.1371/journal.pone.0162226
Cristea, I. A., Mogoaşe, C., David, D., & Cuijpers, P. (2015). Practitioner Review: Cognitive bias modification for mental health problems in children and adolescents: a meta-analysis. Journal of Child Psychology and Psychiatry, 56(7), 723–734. https://doi.org/10.1111/jcpp.12383
Cuijpers, P., Cristea, I. A., Karyotaki, E., Reijnders, M., & Huibers, M. J. H. (2016). How effective are cognitive behavior therapies for major depression and anxiety disorders? A meta-analytic update of the evidence. World Psychiatry, 15(3), 245–258. https://doi.org/10.1002/wps.20346
Cummings, J. J., & Bailenson, J. N. (2016). How Immersive Is Enough? A Meta-Analysis of the Effect of Immersive Technology on User Presence. Media Psychology, 19(2), 272–309. https://doi.org/10.1080/15213269.2015.1015740
Dattani, S., Ritchie, H., & Roser, M. (2021). Mental Health. Our World in Data. https://ourworldindata.org/mental-health
David, D., Matu, S.-A., & David, O. A. (2013). New Directions in Virtual Reality-Based Therapy for Anxiety Disorders. International Journal of Cognitive Therapy, 6(2), 114–137. https://doi.org/10.1521/ijct.2013.6.2.114

Debray, T. P., Schuit, E., Efthimiou, O., Reitsma, J. B., Ioannidis, J. P., Salanti, G., Moons, K. G., & GetReal Workpackage. (2018). An overview of methods for network meta-analysis using individual participant data: When do benefits arise? Statistical Methods in Medical Research, 27(5), 1351–1364. https://doi.org/10.1177/0962280216660741

Dechartres, A., Altman, D. G., Trinquart, L., Boutron, I., & Ravaud, P. (2014). Association between analytic strategy and estimates of treatment outcomes in meta-analyses. JAMA, 312(6), 623–630. https://doi.org/10.1001/jama.2014.8166

Dennis, T. A., & O'Toole, L. (2014). Mental Health on the Go: Effects of a Gamified Attention Bias Modification Mobile Application in Trait Anxious Adults. Clinical Psychological Science : A Journal of the Association for Psychological Science, 2(5), 576–590. https://doi.org/10.1177/2167702614522228

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. Controlled Clinical Trials, 7(3), 177–188. https://doi.org/10.1016/0197-2456(86)90046-2

Derryberry, D., & Reed, M. A. (2002). Anxiety-related attentional biases and their regulation by attentional control. Journal of Abnormal Psychology, 111(2), 225–236. https://doi.org/10.1037/0021-843X.111.2.225

Dias, S., Welton, N. J., Caldwell, D. M., & Ades, A. E. (2010). Checking consistency in mixed treatment comparison meta-analysis. Statistics in Medicine, 29(7–8), 932–944. https://doi.org/10.1002/sim.3767

Dinh, H. Q., Walker, N., Hodges, L. F., Song, C., & Kobayashi, A. (1999). Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. Proceedings IEEE Virtual Reality (Cat. No. 99CB36316), 222–228. https://doi.org/10.1109/VR.1999.756955

Donegan, S., Williamson, P., D'Alessandro, U., & Tudur Smith, C. (2013). Assessing key assumptions of network meta-analysis: A review of methods. Research Synthesis Methods, 4(4), 291–323. https://doi.org/10.1002/jrsm.1085

Dritschel, B. (1992). Cognitive psychology and emotional disorders, J. M. G. Williams, F. Watts, C. Macleod and A. Matthews. Chichester: John Wiley & Sons, 1990. No. of pages 226. ISBN (0-471-929662 (paperback). Price £14.95. Applied Cognitive Psychology, 6(3), 275–277. https://doi.org/10.1002/acp.2350060310

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics, 56(2), 455–463. https://doi.org/10.1111/j.0006-341x.2000.00455.x

Edwards, C. B., Portnow, S., Namaky, N., & Teachman, B. A. (2018). Training Less Threatening Interpretations Over the Internet: Impact of Priming Anxious Imagery and Using a Neutral Control Condition. Cognitive Therapy and Research, 42(6), 832–843. https://doi.org/10.1007/s10608-018-9922-9

Egger, M., Davey Smith, G., Schneider, M. & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. BMJ (Clinical research ed.) 315, 629–634.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. Emotion (Washington, D.C.), 7(2), 336–353. https://doi.org/10.1037/1528-3542.7.2.336

Fodor, L. A., Georgescu, R., Cuijpers, P., Szamoskozi, Ș., David, D., Furukawa, T. A., & Cristea, I. A. (2020). Efficacy of cognitive bias modification interventions in anxiety and depressive disorders: A systematic review and network meta-analysis. The Lancet Psychiatry, 7(6), 506–514. https://doi.org/10.1016/S2215-0366(20)30130-9

Freeman, D., Reeve, S., Robinson, A., Ehlers, A., Clark, D., Spanlang, B., & Slater, M. (2017). Virtual reality in the assessment, understanding, and treatment of mental health disorders. Psychological Medicine, 47(14), 2393–2400. https://doi.org/10.1017/S003329171700040X

Furukawa, T. A. et al. (2014). Waiting list may be a nocebo condition in psychotherapy trials: a contribution from network meta-analysis. Acta psychiatrica Scandinavica 130, 181–192, https://doi.org/10.1111/acps.12275

Furukawa, T. A., & Leucht, S. (2011). How to obtain NNT from Cohen's d: Comparison of two methods. PloS One, 6(4), e19070. https://doi.org/10.1371/journal.pone.0019070

Galbraith, R. F. (1988). A note on graphical presentation of estimated odds ratios from several clinical trials. Statistics in Medicine, 7(8), 889–894. https://doi.org/10.1002/sim.4780070807

Garcia-Palacios, A., Botella, C., Hoffman, H., & Fabregat, S. (2007). Comparing Acceptance and Refusal Rates of Virtual Reality Exposure vs. In Vivo Exposure by Patients with Specific Phobias. CyberPsychology & Behavior, 10(5), 722–724. https://doi.org/10.1089/cpb.2007.9962

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. https://doi.org/10.1007/s11920-010-0128-4

Gorini, A., & Riva, G. (2008). Virtual reality in anxiety disorders: The past and the future. Expert Review of Neurotherapeutics, 8(2), 215–233. https://doi.org/10.1586/14737175.8.2.215

Greenland, S., & Robins, J. M. (1985). Estimation of a common effect parameter from sparse follow-up data. Biometrics, 41(1), 55–68.

Grisham, J. R., Becker, L., Williams, A. D., Whitton, A. E., & Makkar, S. R. (2014). Using cognitive bias modification to deflate responsibility in compulsive checkers. Cognitive Therapy and Research, 38(5), 505–517. https://doi.org/10.1007/s10608-014-9621-0

Hallion, L. S., & Ruscio, A. M. (2011). A meta-analysis of the effect of cognitive bias modification on anxiety and depression. Psychological Bulletin, 137(6), 940–958. https://doi.org/10.1037/a0024355

Harbord, R. M., Egger, M., & Sterne, J. A. C. (2006). A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. Statistics in Medicine, 25(20), 3443–3457. https://doi.org/10.1002/sim.2380

Harris, S. R., Kemmerling, R. L., & North, M. M. (2002). Brief Virtual Reality Therapy for Public Speaking Anxiety. CyberPsychology & Behavior, 5(6), 543–550. https://doi.org/10.1089/109493102321018187

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Eds.), Advances in Psychology (Vol. 52, pp. 139–183). North-Holland. https://doi.org/10.1016/S0166-4115(08)62386-9

Hedges, L. V., Olkin, I., & Hedges, L. V. (1985). Statistical Methods for Meta-Analysis. https://www.scholars.northwestern.edu/en/publications/statistical-methods-for-meta-analysis

Heeren, A., Mogoaşe, C., Philippot, P., & McNally, R. J. (2015). Attention bias modification for social anxiety: A systematic review and meta-analysis. Clinical Psychology Review, 40, 76–90. https://doi.org/10.1016/j.cpr.2015.06.001

Henricks, L.A., Lange, W.G., Luijten, M. et al. (2022). The Longitudinal Interplay Between Attention Bias and Interpretation Bias in Social Anxiety in Adolescents. Cogn Ther Res (2022). https://doi.org/10.1007/s10608-022-10304-1

Higgins, I., & Green, S. (2011). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. http://training.cochrane.org/handbook

Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., Savović, J., Schulz, K. F., Weeks, L., & Sterne, J. A. C. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ, 343, d5928. https://doi.org/10.1136/bmj.d5928

Higgins, J. P. T., Jackson, D., Barrett, J. K., Lu, G., Ades, A. E., & White, I. R. (2012). Consistency and inconsistency in network meta-analysis: Concepts and models for multi-arm studies. Research Synthesis Methods, 3(2), 98–110. https://doi.org/10.1002/jrsm.1044

Higgins, J. P. T., Thompson, S. G., & Spiegelhalter, D. J. (2009). A re-evaluation of random-effects meta-analysis. Journal of the Royal Statistical Society. Series A, (Statistics in Society), 172(1), 137–159. https://doi.org/10.1111/j.1467-985X.2008.00552.x

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. BMJ, 327(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

Higgins, J. P., Sterne, J. A., Savovic, J., Page, M. J., Hróbjartsson, A., Boutron, I., Reeves, B., & Eldridge, S. (2016). A revised tool for assessing risk of bias in randomized trials. Cochrane Database of Systematic Reviews, 10(Suppl 1), 29–31.

Hirschfeld, R. M. A. (2001). The Comorbidity of Major Depression and Anxiety Disorders: Recognition and Management in Primary Care. Primary Care Companion to The Journal of Clinical Psychiatry, 3(6), 244–254.

Hodges, L. F., Watson, B. A., Kessler, G. D., Rothbaum, B. O., & Opdyke, D. (1996). Virtually conquering fear of flying. IEEE Computer Graphics and Applications, 16(6), 42–49. Scopus. https://doi.org/10.1109/38.544071

Holmes, E. A., Ghaderi, A., Harmer, C. J., Ramchandani, P. G., Cuijpers, P., Morrison, A. P., Roiser, J. P., Bockting, C. L. H., O'Connor, R. C., Shafran, R., Moulds, M. L., & Craske, M. G. (2018). The Lancet Psychiatry Commission on psychological treatments research in tomorrow's science. The Lancet Psychiatry, 5(3), 237–286. https://doi.org/10.1016/S2215-0366(17)30513-8

Hutton, B., Salanti, G., Caldwell, D. M., Chaimani, A., Schmid, C. H., Cameron, C., Ioannidis, J. P. A., Straus, S., Thorlund, K., Jansen, J. P., Mulrow, C., Catalá-López, F., Gøtzsche, P. C., Dickersin, K., Boutron, I., Altman, D.

G., & Moher, D. (2015). The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: Checklist and explanations. Annals of Internal Medicine, 162(11), 777–784. https://doi.org/10.7326/M14-2385

IntHout, J., Ioannidis, J. P. A., Borm, G. F., & Goeman, J. J. (2015). Small studies are more heterogeneous than large ones: A meta-meta-analysis. Journal of Clinical Epidemiology, 68(8), 860–869. https://doi.org/10.1016/j.jclinepi.2015.03.017

Ioannidis, J. P. (1998). Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. JAMA, 279(4), 281–286. https://doi.org/10.1001/jama.279.4.281

Ioannidis, J. P. A., Patsopoulos, N. A., & Evangelou, E. (2007). Uncertainty in heterogeneity estimates in meta-analyses. BMJ, 335(7626), 914–916. https://doi.org/10.1136/bmj.39343.408449.80

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. Cognitive Psychology, 3(3), 430–454. https://doi.org/10.1016/0010-0285(72)90016-3

Kalin, N. H. (2020). The Critical Relationship Between Anxiety and Depression. American Journal of Psychiatry, 177(5), 365–367. https://doi.org/10.1176/appi.ajp.2020.20030305

Karoly, P., & Crombez, G. (2018). Motivational Perspectives on Chronic Pain: Theory, Research, and Practice. Oxford University Press.

Kaussner, Y., Kuraszkiewicz, A. M., Schoch, S., Markel, P., Hoffmann, S., Baur-Streubel, R., Kenntner-Mabiala, R., & Pauli, P. (2020). Treating patients with driving phobia by virtual reality exposure therapy – a pilot study. PLOS ONE, 15(1), e0226937. https://doi.org/10.1371/journal.pone.0226937

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. The International Journal of Aviation Psychology, 3(3), 203–220. https://doi.org/10.1207/s15327108ijap0303_3

Kessler, R. C., Stang, P. E., Wittchen, H. U., Ustun, T. B., Roy-Burne, P. P., & Walters, E. E. (1998). Lifetime panic-depression comorbidity in the National Comorbidity Survey. Archives of General Psychiatry, 55(9), 801–808. https://doi.org/10.1001/archpsyc.55.9.801

Klinger, E., Bouchard, S., Légeron, P., Roy, S., Lauer, F., Chemin, I., & Nugues, P. (2005). Virtual Reality Therapy Versus Cognitive Behavior Therapy for Social Phobia: A Preliminary Controlled Study. CyberPsychology & Behavior, 8(1), 76–88. https://doi.org/10.1089/cpb.2005.8.76

Koster, E. H. W., Crombez, G., Verschuere, B., & De Houwer, J. (2004). Selective attention to threat in the dot probe paradigm: Differentiating vigilance and difficulty to disengage. Behaviour Research and Therapy, 42(10), 1183–1192. https://doi.org/10.1016/j.brat.2003.08.001

Kraemer, H. C., & Kupfer, D. J. (2006). Size of Treatment Effects and Their Importance to Clinical Research and Practice. Biological Psychiatry, 59(11), 990–996. https://doi.org/10.1016/j.biopsych.2005.09.014

Krings, A., Heeren, A., Fontaine, P., & Blairy, S. (2020). Attentional biases in depression: Relation to disorder severity, rumination, and anhedonia.

Comprehensive psychiatry, 100, 152173. https://doi.org/10.1016/j.comppsych.2020.152173

Lamers, F., van Oppen, P., Comijs, H. C., Smit, J. H., Spinhoven, P., van Balkom, A. J. L. M., Nolen, W. A., Zitman, F. G., Beekman, A. T. F., & Penninx, B. W. J. H. (2011). Comorbidity patterns of anxiety and depressive disorders in a large cohort study: The Netherlands Study of Depression and Anxiety (NESDA). The Journal of Clinical Psychiatry, 72(3), 341–348. https://doi.org/10.4088/JCP.10m06176blu

Laupacis, A., Sackett, D. L., & Roberts, R. S. (1988). An Assessment of Clinically Useful Measures of the Consequences of Treatment. New England Journal of Medicine, 318(26), 1728–1733. https://doi.org/10.1056/NEJM198806303182605

Leary, M. R. (1983). A Brief Version of the Fear of Negative Evaluation Scale. Personality and Social Psychology Bulletin, 9(3), 371–375. https://doi.org/10.1177/0146167283093007

Leucht, S., Chaimani, A., Cipriani, A. S., Davis, J. M., Furukawa, T. A., & Salanti, G. (2016). Network meta-analyses should be the highest level of evidence in treatment guidelines. European Archives of Psychiatry and Clinical Neuroscience, 266(6), 477–480. https://doi.org/10.1007/s00406-016-0715-4

Linetzky, M., Pergamin-Hight, L., Pine, D. S., & Bar-Haim, Y. (2015). Quantitative evaluation of the clinical efficacy of attention bias modification treatment for anxiety disorders. Depression and Anxiety, 32(6), 383–391. https://doi.org/10.1002/da.22344

Liu, H., Li, X., Han, B., & Liu, X. (2017). Effects of cognitive bias modification on social anxiety: A meta-analysis. PloS One, 12(4), e0175107. https://doi.org/10.1371/journal.pone.0175107

Long, J. S., & Long, J. S. (1997). Regression Models for Categorical and Limited Dependent Variables. SAGE.

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska directed emotional faces (KDEF). CD ROM from Department of Clinical Neuroscience, Psychology Section, Karolinska Institutet, 91(630), 2–2.

Ma, L., Kruijt, A.-W., Nöjd, S., Zetterlund, E., Andersson, G., & Carlbring, P. (2019). Attentional Bias Modification in Virtual Reality – A VR-Based Dot-Probe Task With 2D and 3D Stimuli. Frontiers in Psychology, 10, 2526. https://doi.org/10.3389/fpsyg.2019.02526

MacDonald, E. M., Koerner, N., & Antony, M. M. (2013). Modification of Interpretive Bias: Impact on Anxiety Sensitivity, Information Processing and Response to Induced Bodily Sensations. Cognitive Therapy and Research, 37(4), 860–871. https://doi.org/10.1007/s10608-012-9519-7

MacLeod, C., & Mathews, A. (2012). Cognitive bias modification approaches to anxiety. Annual Review of Clinical Psychology, 8, 189–217. https://doi.org/10.1146/annurev-clinpsy-032511-143052

MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. Journal of Abnormal Psychology, 95(1), 15–20. https://doi.org/10.1037/0021-843X.95.1.15

Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. Journal of the National Cancer Institute, 22(4), 719–748.

McCann, R. A., Armstrong, C. M., Skopp, N. A., Edwards-Stewart, A., Smolenski, D. J., June, J. D., Metzger-Abamukong, M., & 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. https://doi.org/10.1016/j.janxdis.2014.05.010

McLay, R. N., Baird, A., Webb-Murphy, J., Deal, W., Tran, L., Anson, H., Klam, W., & Johnston, S. (2017). A Randomized, Head-to-Head Study of Virtual Reality Exposure Therapy for Posttraumatic Stress Disorder. Cyberpsychology, Behavior, and Social Networking, 20(4), 218–224. https://doi.org/10.1089/cyber.2016.0554

Menne-Lothmann, C., Viechtbauer, W., Höhn, P., Kasanova, Z., Haller, S. P., Drukker, M., van Os, J., Wichers, M., & Lau, J. Y. F. (2014). How to boost positive interpretations? A meta-analysis of the effectiveness of cognitive bias modification for interpretation. PloS One, 9(6), e100925. https://doi.org/10.1371/journal.pone.0100925

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. https://doi.org/10.1002/da.20734

Miloff, A., Lindner, P., Hamilton, W., Reuterskiöld, L., Andersson, G., & Carlbring, P. (2016). Single-session gamified virtual reality exposure therapy for spider phobia vs. traditional exposure therapy: Study protocol for a randomized controlled non-inferiority trial. Trials, 17(1), 60. https://doi.org/10.1186/s13063-016-1171-1

Mogg, K., & Bradley, B. P. (1998). A cognitive-motivational analysis of anxiety. Behaviour Research and Therapy, 36(9), 809–848. https://doi.org/10.1016/S0005-7967(98)00063-1

Mogg, K., Bradley, B. P., & Williams, R. (1995). Attentional bias in anxiety and depression: The role of awareness. British Journal of Clinical Psychology, 34(1), 17–36. https://doi.org/10.1111/j.2044-8260.1995.tb01434.x

Mogg, K., McNamara, J., Powys, M., Rawlinson, H., Seiffer, A., & Bradley, B. P. (2000). Selective attention to threat: A test of two cognitive models of anxiety. Cognition and Emotion, 14(3), 375–399. https://doi.org/10.1080/026999300378888

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. BMJ, 339, b2535. https://doi.org/10.1136/bmj.b2535

North, M. M., & Rives, J. (2003). Virtual reality therapy in aid of senior citizens' psychological disorders. Studies in Health Technology and Informatics, 94, 245–247.

North, M. M., North, S. M., & Coble, J. R. (1995). Effectiveness of Virtual Environment Desensitization in the Treatment of Agoraphobia. International Journal of Virtual Reality, 1(2), 25–34. https://doi.org/10.20870/IJVR.1995.1.2.2603

North, M. M., North, S. M., & Coble, J. R. (1996a). Virtual reality therapy: An innovative paradigm. IPI Press. http://books.google.com/books?id=nHlHAAAAMAAJ

North, M. M., North, S. M., & Coble, J. R. (1996b). Effectiveness of Virtual Environment Desensitization in the Treatment of Agoraphobia. Presence:

Teleoperators and Virtual Environments, 5(3), 346–352. https://doi.org/10.1162/pres.1996.5.3.346
North, M. N., & North, S. M. (2000). Virtual reality combats obsessive-compulsive disorders (OCD). Proceedings of Medicine Meets Virtual Reality.
Nowakowski, M. E., Antony, M. M., & Koerner, N. (2015). Modifying interpretation biases: Effects on symptomatology, behavior, and physiological reactivity in social anxiety. Journal of Behavior Therapy and Experimental Psychiatry, 49(Pt A), 44–52. https://doi.org/10.1016/j.jbtep.2015.04.004
Nüesch, E., Trelle, S., Reichenbach, S., Rutjes, A. W. S., Bürgi, E., Scherer, M., Altman, D. G., & Jüni, P. (2009). The effects of excluding patients from the analysis in randomised controlled trials: Meta-epidemiological study. BMJ, 339, b3244. https://doi.org/10.1136/bmj.b3244
Nüesch, E., Trelle, S., Reichenbach, S., Rutjes, A. W. S., Tschannen, B., Altman, D. G., Egger, M., & Jüni, P. (2010). Small study effects in meta-analyses of osteoarthritis trials: Meta-epidemiological study. BMJ (Clinical Research Ed.), 341, c3515. https://doi.org/10.1136/bmj.c3515
Olfson, M., Fireman, B., Weissman, M. M., Leon, A. C., Sheehan, D. V., Kathol, R. G., Hoven, C., & Farber, L. (1997). Mental disorders and disability among patients in a primary care group practice. The American Journal of Psychiatry, 154(12), 1734–1740. https://doi.org/10.1176/ajp.154.12.1734
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. https://doi.org/10.1002/da.20910
Orsini, N., Bottai, M., Higgins, J., & Buchan, I. (2006). HETEROGI: Stata module to quantify heterogeneity in a meta-analysis. https://econpapers.repec.org/software/bocbocode/s449201.htm
Pacheco-Unguetti, A. P., Acosta, A., Marqués, E., & Lupiáñez, J. (2011). Alterations of the attentional networks in patients with anxiety disorders. Journal of Anxiety Disorders, 25(7), 888–895. https://doi.org/10.1016/j.janxdis.2011.04.010
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. https://doi.org/10.1016/j.jbtep.2007.07.007
Peckham, A. D., McHugh, R. K., & Otto, M. W. (2010). A meta-analysis of the magnitude of biased attention in depression. Depression and anxiety, 27(12), 1135–1142. https://doi.org/10.1002/da.20755
Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. Journal of Clinical Epidemiology, 61(10), 991–996. https://doi.org/10.1016/j.jclinepi.2007.11.010
Phillips, M. L., Senior, C., Fahy, T., & David, A. S. (1998). Disgust – the forgotten emotion of psychiatry. The British Journal of Psychiatry, 172(5), 373–375. https://doi.org/10.1192/bjp.172.5.373
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. https://doi.org/10.1016/j.janxdis.2007.04.006

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. https://doi.org/10.1016/j.janxdis.2011.03.004

Price, R. B., Wallace, M., Kuckertz, J. M., Amir, N., Graur, S., Cummings, L., Popa, P., Carlbring, P., & Bar-Haim, Y. (2016). Pooled patient-level meta-analysis of children and adults completing a computer-based anxiety intervention targeting attentional bias. Clinical Psychology Review, 50, 37–49. https://doi.org/10.1016/j.cpr.2016.09.009

R Core, T. (2018). R: A language and environment for statistical computing [R; R]. R Foundation for Statistical Computing. https://www.R-project.org/.

Raftery, A. E. (1995). Bayesian Model Selection in Social Research. Sociological Methodology, 25, 111–163. https://doi.org/10.2307/271063

Rausch, J. R., Maxwell, S. E., & Kelley, K. (2003). Analytic Methods for Questions Pertaining to a Randomized Pretest, Posttest, Follow-Up Design. Journal of Clinical Child & Adolescent Psychology, 32(3), 467–486. https://doi.org/10.1207/S15374424JCCP3203_15

Reger, G. M., Koenen-Woods, P., Zetocha, K., Smolenski, D. J., Holloway, K. M., Rothbaum, B. O., Difede, J., Rizzo, A. A., Edwards-Stewart, A., Skopp, N. A., Mishkind, M., Reger, M. A., & Gahm, G. A. (2016). Randomized controlled trial of prolonged exposure using imaginal exposure vs. Virtual reality exposure in active duty soldiers with deployment-related posttraumatic stress disorder (PTSD). Journal of Consulting and Clinical Psychology, 84(11), 946–959. https://doi.org/10.1037/ccp0000134

Riva, G. (1997). The Virtual Environment for Body-Image Modification (VEBIM): Development and Preliminary Evaluation. Presence: Teleoperators and Virtual Environments, 6(1), 106–117. https://doi.org/10.1162/pres.1997.6.1.106

Rizzo, A. “Skip,” Difede, J., Rothbaum, B. O., Reger, G., Spitalnick, J., Cukor, J., & Mclay, R. (2010). Development and early evaluation of the Virtual Iraq/Afghanistan exposure therapy system for combat-related PTSD. Annals of the New York Academy of Sciences, 1208(1), 114–125. https://doi.org/10.1111/j.1749-6632.2010.05755.x

Rizzo, A. a., Buckwalter, J. g., Bowerly, T., Van Der Zaag, C., Humphrey, L., Neumann, U., Chua, C., Kyriakakis, C., Van Rooyen, A., & Sisemore, D. (2000). The Virtual Classroom: A Virtual Reality Environment for the Assessment and Rehabilitation of Attention Deficits. CyberPsychology & Behavior, 3(3), 483–499. https://doi.org/10.1089/10949310050078940

Robbins, T. W., Weinberger, D., Taylor, J. G., Morris, R. G., Roberts, A. C., Robbins, T. W., & Weiskrantz, L. (1996). Dissociating executive functions of the prefrontal cortex. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 351(1346), 1463–1471. https://doi.org/10.1098/rstb.1996.0131

Rock, P. L., Roiser, J. P., Riedel, W. J., & Blackwell, A. D. (2014). Cognitive impairment in depression: A systematic review and meta-analysis. Psychological Medicine, 44(10), 2029–2040. https://doi.org/10.1017/S0033291713002535

Rothbaum, B. O., Hodges, L. F., Kooper, R., Opdyke, D., Williford, J. S., & North, M. (1995a). Effectiveness of computer-generated (virtual reality) graded exposure in the treatment of acrophobia. American Journal of Psychiatry, 152(4), 626–628. Scopus. https://doi.org/10.1176/ajp.152.4.626

Rothbaum, B. O., Hodges, L. F., Kooper, R., Opdyke, D., Williford, J. S., & North, M. (1995b). Virtual reality graded exposure in the treatment of acrophobia: A case report. Behavior Therapy, 26(3), 547–554. https://doi.org/10.1016/S0005-7894(05)80100-5

Rothbaum, B. O., Hodges, L., Alarcon, R., Ready, D., Shahar, F., Graap, K., Pair, J., Hebert, P., Gotz, D., Wills, B., & Baltzell, D. (1999). Virtual Reality Exposure Therapy for PTSD Vietnam Veterans: A Case Study. Journal of Traumatic Stress, 12(2), 263–271. https://doi.org/10.1023/A:1024772308758

Rucker, G., Krahn, U., Konig, J., Efthimiou, O., & Schwaiger, G. (2019). Network Meta-Analysis using Frequentist Methods. https://github.com/guido-s/netmeta http://meta-analysis-with-r.org

Sakaluk, J. K., Williams, A. J., Kilshaw, R. E., & Rhyner, K. T. (2019). Evaluating the evidential value of empirically supported psychological treatments (ESTs): A meta-scientific review. Journal of Abnormal Psychology, 128(6), 500–509. https://doi.org/10.1037/abn0000421

Sass, S. M., Evans, T. C., Xiong, K., Mirghassemi, F., & Tran, H. (2017). Attention training to pleasant stimuli in anxiety. Biological Psychology, 122, 80–92. https://doi.org/10.1016/j.biopsycho.2016.03.003

Savović, J., Turner, R. M., Mawdsley, D., Jones, H. E., Beynon, R., Higgins, J. P. T., & Sterne, J. A. C. (2018). Association Between Risk-of-Bias Assessments and Results of Randomized Trials in Cochrane Reviews: The ROBES Meta-Epidemiologic Study. American Journal of Epidemiology, 187(5), 1113–1122. https://doi.org/10.1093/aje/kwx344

Scammacca, N., Roberts, G., & Stuebing, K. K. (2014). Meta-Analysis With Complex Research Designs: Dealing With Dependence From Multiple Measures and Multiple Group Comparisons. Review of Educational Research, 84(3), 328–364. https://doi.org/10.3102/0034654313500826

Schumacher, S., Betzler, F., Miller, R., Kirschbaum, C., & Ströhle, A. (2017). Habituation of stress in psychotherapists performing subsequent in vivo exposures—A case series. Journal of Psychotherapy Integration, 27(2), 218–224. https://doi.org/10.1037/int0000086

Sherbourne, C. D., Wells, K. B., Meredith, L. S., Jackson, C. A., & Camp, P. (1996). Comorbid anxiety disorder and the functioning and well-being of chronically ill patients of general medical providers. Archives of General Psychiatry, 53(10), 889–895. https://doi.org/10.1001/archpsyc.1996.01830100035005

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments five: Speculations on the role of presence in virtual environments. Presence: Teleoperators and Virtual Environments, 6(6), 603–616. https://doi.org/10.1162/pres.1997.6.6.603

Smeeth, L., Haines, A., & Ebrahim, S. (1999). Numbers needed to treat derived from meta-analyses—Sometimes informative, usually misleading. BMJ, 318(7197), 1548–1551. https://doi.org/10.1136/bmj.318.7197.1548

Spielberger, C. D., Gorssuch, R. L., Lushene, P. R., Vagg, P. R., & Jacobs, G. A. (1983). Manual for the State-Trait Anxiety Inventory. Consulting Psychologists Press.
Spielberger, C. D., Pitariu, H., & Pleasa, C. (2007). STAI-Y. State-trait anxiety inventory. Sinapsis.
StataCorp.2017. (2017). Stata Statistical Software: Release 15. (Version 15) [Computer software]. College Station.
Steinman, S. A., & Teachman, B. A. (2010). Modifying interpretations among individuals high in anxiety sensitivity. Journal of Anxiety Disorders, 24(1), 71–78. https://doi.org/10.1016/j.janxdis.2009.08.008
Sterne, J. A., Gavaghan, D., & Egger, M. (2000). Publication and related bias in meta-analysis: Power of statistical tests and prevalence in the literature. Journal of Clinical Epidemiology, 53(11), 1119–1129. https://doi.org/10.1016/s0895-4356(00)00242-0
Strickland, D. (1996). A Virtual Reality Application with Autistic Children. Presence: Teleoperators and Virtual Environments, 5(3), 319–329. https://doi.org/10.1162/pres.1996.5.3.319
Stuss, D. T., & Alexander, M. P. (2000). Executive functions and the frontal lobes: A conceptual view. Psychological Research, 63(3), 289–298. https://doi.org/10.1007/s004269900007
Susa, G., Pitică, I., Benga, O., & Miclea, M. (2012). The self regulatory effect of attentional control in modulating the relationship between attentional biases toward threat and anxiety symptoms in children. Cognition and Emotion, 26(6), 1069–1083. https://doi.org/10.1080/02699931.2011.638910
Sutherland, I. (1965). The Ultimate Display. Proceedings of the IFIPS Congress 65(2):506-508. New York: IFIP. http://papers.cumincad.org/cgi-bin/works/Show&_id=caadria2010_039/paper/c58e
Turner, W. A., & Casey, L. M. (2014). Outcomes associated with virtual reality in psychological interventions: Where are we now? Clinical Psychology Review, 34(8), 634–644. https://doi.org/10.1016/j.cpr.2014.10.003
Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. Science, 185(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124
Twenge, J. M., & Joiner, T. E. (2020). U.S. Census Bureau-assessed prevalence of anxiety and depressive symptoms in 2019 and during the 2020 COVID-19 pandemic. Depression and Anxiety. https://doi.org/10.1002/da.23077
Urech, A., Krieger, T., Chesham, A., Mast, F. W., & Berger, T. (2015). Virtual Reality-Based Attention Bias Modification Training for Social Anxiety: A Feasibility and Proof of Concept Study. Frontiers in Psychiatry, 6, 154. https://doi.org/10.3389/fpsyt.2015.00154
Van Ryckeghem, D. M. L., & Crombez, G. (2014). Attentional bias and chronic pain: Where to go from here? Pain, 155(1), 6–7. https://doi.org/10.1016/j.pain.2013.10.025
Van Ryckeghem, D. M. L., Van Damme, S., & Vervoort, T. (2018). Does attention bias modification training impact on task performance in the context of pain: An experimental study in healthy participants. PLoS ONE, 13(7), e0200629. https://doi.org/10.1371/journal.pone.0200629
Vermeulen, M., Brown, A. D., Raes, F., & Krans, J. (2019). Decreasing event centrality in undergraduates using cognitive bias modification of appraisals.

Cognitive Therapy and Research, 43(1), 214–225. https://doi.org/10.1007/s10608-018-9936-3
Veroniki, A. A., Vasiliadis, H. S., Higgins, J. P. T., & Salanti, G. (2013). Evaluation of inconsistency in networks of interventions. International Journal of Epidemiology, 42(1), 332–345. https://doi.org/10.1093/ije/dys222
Wallach, H., Safir, M., Horef, R., Huber, E., & Heiman, T. (2012). Presence in virtual reality: Importance and methods to increase it. Virtual Reality, 107–123.
White, I. (2015). Network meta-analysis. The Stata Journal, 15(4), 951–985.
White, I. R. (2011). Multivariate Random-effects Meta-regression: Updates to Mvmeta. The Stata Journal, 11(2), 255–270. https://doi.org/10.1177/1536867X1101100206
White, I. R., Barrett, J. K., Jackson, D., & Higgins, J. P. T. (2012). Consistency and inconsistency in network meta-analysis: Model estimation using multivariate meta-regression. Research Synthesis Methods, 3(2), 111–125. https://doi.org/10.1002/jrsm.1045
Wiederhold, B. K., Gevirtz, R., & Wiederhold, M. D. (1998). Fear of Flying: A Case Report Using Virtual Reality Therapy with Physiological Monitoring. CyberPsychology & Behavior, 1(2), 97–103. https://doi.org/10.1089/cpb.1998.1.97
Williams, J. M. G., Watts, F. N., Macleod, C., & Mathews, A. (1997). Cognitive Psychology and Emotional Disorders. John Wiley & Sons.
Wilson, E., & MacLeod, C. (2003). Contrasting two accounts of anxiety-linked attentional bias: Selective attention to varying levels of stimulus threat intensity. Journal of Abnormal Psychology, 112(2), 212–218. https://doi.org/10.1037/0021-843x.112.2.212
Witmer, B. G., & Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. Presence: Teleoperators and Virtual Environments, 7(3), 225–240. https://doi.org/10.1162/105474698565686
Wood, D. P., Murphy, J., McLay, R., Koffman, R., Spira, J., Obrecht, R. E., Pyne, J., & Wiederhold, B. K. (2009). Cost Effectiveness of Virtual Reality Graded Exposure Therapy with Physiological Monitoring for the Treatment of Combat Related Post Traumatic Stress Disorder. Annual Review of Cybertherapy and Telemedicine 2009, 223–229. https://doi.org/10.3233/978-1-60750-017-9-223
Yusuf, S., Peto, R., Lewis, J., Collins, R., & Sleight, P. (1985). Beta blockade during and after myocardial infarction: An overview of the randomized trials. Progress in Cardiovascular Diseases, 27(5), 335–371. https://doi.org/10.1016/s0033-0620(85)80003-7
Zvielli A., Bernstein A., Koster E.H.W. (2014). Dynamics of Attentional Bias to Threat in Anxious Adults: Bias towards and/or Away? PLoS ONE 9(8): e104025. https://doi.org/10.1371/journal.pone.0104025