Appendix A

Description of interventions and criteria for classification

category	description of the intervention	Classification criteria
EMDR	Eye movement desensitization and reprocessing (EMDR) is a psychotherapy treatment based on Shapiro's (2001) Adaptive Information Processing model which postulates that EMDR therapy facilitates the accessing and processing of traumatic memories and other adverse life events. During EMDR therapy the client attends to the flow of associations regarding emotionally disturbing material in brief sequential doses. Simultaneously, the client is asked to follow the therapists index finger moving back and forth about twelve inches from the client's eyes. Therapist directed lateral eye movements represent the most commonly used external stimulus, but other stimuli can be used including hand-tapping and audio stimulation. Shapiro (2001) hypothesizes that EMDR therapy facilitates the accessing of the traumatic memory network, so that information processing is enhanced, with new associations forged between the traumatic memory and more adaptive memories or information. These new associations are thought to result in complete information processing, new learning, elimination of emotional distress, and development of cognitive insights. Recent research suggests a link between EMDR and REM-sleep, proposing that Alternate Bilateral Stimulations typical of EMDR shift the brain into a memory processing mode similar to that of REM sleep (e.g., Pagani & Carletto, 2017).	Code if participants received manualized individual eye movement desensitization and reprocessing with visual, auditory and/or tactile stimulation. Do not code if no eye movements were included or if only a single EMDR session was implemented.
PE	Prolonged exposure (PE) is recommended as a first-line treatment approach for PTSD (e.g., NICE, 2018). PE focuses on reducing the intense negative emotions that are caused by memories or being reminded of the trauma. The theorized mechanisms underlying PE are based on emotional processing theory and broader extinction models of fear reduction. Core components are breathing retraining, imaginal re-experiencing, and in vivo exposure. As between-session homework, patients are also instructed to listen daily to a recording of their imaginal exposure. PE is typically concluded after eight to fifteen 90-minute sessions.	Code if intervention follows a PE treatment manual. Core components: breathing retraining, imaginal re-experiencing, and in vivo exposure. As between-session homework, patients listen daily to a recording of their imaginal exposure. Do not code if PE is combined with another intervention, if

CPT	Cognitive Processing Therapy (CPT) is a 12-session therapy that has been found effective for PTSD and other corollary symptoms following traumatic events (Monson et al., 2006; Resick et al., 2002). CPT is based on a social cognitive theory of PTSD that focuses on how the traumatic event is coped with by a person who is trying to regain a sense of mastery and control in his or her life. The other major theory explaining PTSD is Lang's (1977) information processing theory, which was extended to PTSD by Foa, Steketee, and Rothbaum (1989) in their emotional processing theory of PTSD. Changing the content of cognitions about a trauma can impact emotional and behavioral responses to the trauma. In doing so, the patient creates a new understanding and conceptualization of the traumatic event so that it reduces its ongoing negative effects on current life. After psychoeducation the client writes an impact statement that details current understanding of why the traumatic event occurred and the impact it has had on beliefs about self, others, and the world. Then the client begins more formal processing of the trauma by writing a detailed account of the worst traumatic experience. Automatic thoughts and beliefs are identified and modified using cognitive techniques. The client develops skills to identify and address unhelpful thinking.	only a single session was implemented, or if the intervention was not delivered in an individual format. Code if manualized multi-session individual CPT was implemented, including core components: psychoeducation, automatic thoughts analysis, writing an impact statement, and an account of the worst traumatic experience; using cognitive techniques.
CT	Cognitive Therapy for PTSD (CT-PTSD) is based on Ehlers and Clark's (2000) model of PTSD. It focuses on three factors specified in this model. According to this theory, people suffering from PTSD perceive a serious current threat which has two sources, excessively negative appraisals of the trauma and/or its sequelae and characteristics of trauma memories that lead to re-experiencing symptoms. The problem is maintained by cognitive strategies and behaviors (such as thought suppression, rumination, and safety-seeking behaviors) that are intended to reduce the sense of current threat but maintain the problem by preventing change in the appraisals or trauma memory, and/or by increasing symptoms (Ehlers et al., 2013). The therapist uses a strategy of Socratic questioning along with other approaches to help the patient arrive at a different evaluation of the traumatic event. This new evaluation is then integrated with the traumatic memory or cues. This can be	Code if cognitive techniques were implemented targeting PTSD (e.g. cognitive restructuring, CT for PTSD) without exposure which aims at habituation. Do not code if only a single CT session was implemented or if not delivered in an individual format.

accomplished by the patient writing and thinking about the new evaluation while at the same time considering the trauma memory, or by embedding the new evaluation into a subsequent imaginal reliving of the traumatic experience. Although this can have an 'element' of exposure, the intention is to identify idiosyncratic memory points eliciting strong responses and to use cognitive restructuring in the moment.

NET	Narrative exposure therapy (NET) has been developed as a short and pragmatic treatment approach (manualized in Schauer, Neuner, & Elbert, 2011) for application in low-income countries affected by war and human rights violations. Within four to 14 individual 90-minute sessions, the client and therapist create a written autobiography containing the major emotional memories of the survivor from birth to the present. The focus of NET is on reconstructing the fragmented memories of traumatic experiences into coherent narrations that are connected to the temporal and spatial context of the life period. At the end of treatment, a copy of the final consistent life narration is handed over to the client, and the therapist keeps one copy that may, depending on the wishes of the client, be used for human rights purposes. NET has been developed based on the following principles: trauma-focus, life-span approach, advocacy (recognition of the victim's suffering, also in human rights context), task-shifting (application by trained local lay health workers) and a cross-cultural approach. The rationale of NET is based on current psychological theories of PTSD that commonly identify dysfunctional memory processes as being at the core of the disorder (Ehlers & Clark, 2000).	Code if manualized individual NET was implemented, where client and therapist create a written autobiography containing major emotional memories from birth to the present. Do not code if only a single session was administered.
SIT	Stress Inoculation Therapy (SIT) is a psychotherapy method intended to help patients prepare themselves in advance to handle stressful events successfully and with a minimum of upset. SIT has three phases: In the initial conceptualization phase, the therapist educates the patient about the general nature of stress and works to develop a clear understanding of the nature of the stressors the patient is facing. The second phase of SIT focuses on skills acquisition and rehearsal. In the final SIT phase, application and follow through, the therapist provides the patient with opportunities to practice coping skills. In most instances, SIT consists of eight to fifteen sessions, plus booster and follow-up sessions, conducted over a three-to-twelve-month period.	Code if manualized SIT in an individual format was implemented, which focuses on skills acquisition and rehearsal. Other types of stress management training are not coded here. Do not code if only a single session of SIT was delivered.

IPT	Interpersonal psychotherapy (IPT) is a time-limited, structured, manualized, evidence-based approach to treat mood disorders. The main goal of IPT is to improve the quality of a client's interpersonal relationships (in the here and now) and social functioning to help reduce their distress. IPT provides strategies to resolve problems within four key areas: role disputes, role transitions, unresolved grief, and interpersonal deficits. Length of treatment is usually 12 to 16 weeks.	Code if individual manualized multi-session IPT was implemented, focussing on the client's interpersonal relationships.
PCT	Present-centered therapy (PCT) was originally developed as a strong comparator treatment that captured many of the effective components of "good psychotherapy" to test whether trauma-focused cognitive-behavioral therapy (TF-CBT) demonstrated effects beyond common psychotherapeutic benefits (Schnurr et al., 2003). Components of PCT include (1) psychoeducation on PTSD, (2) strategies for approaching day-to-day challenges; and (3) homework outside the session. PCT is often described by stating what the treatment does not entail: Therapy is not trauma-focused (i.e., PCT does not include disclosure, discussion, or exposure of traumatic events); therapy is not based on a cognitive-behavioral therapy framework (i.e., PCT does not focus on cognitive restructuring or graded exposure); and therapy is not strictly supportive (i.e., PCT is a structured treatment with homework assigned between sessions). PCT typically is modified to mirror the active treatment under investigation and can be delivered in both individual and group formats, with length of treatment and duration of sessions based on the active treatment.	Code if individual manualized PCT was implemented, including psychoeducation on PTSD, strategies for approaching day-to-day challenges, and homework outside the session. Do not code if any disclosure, discussion or exposure regarding the traumatic event was implemented or if only a single PCT session was implemented.
MBI	Mindfulness meditation has a longstanding history in eastern practices that has received considerable public interest in recent decades. The science, practice, and implementation of mindfulness-based interventions (MBIs) have dramatically increased in recent years. Mindfulness is defined as "paying attention in a particular way, on purpose, in the present moment, and nonjudgmentally" (Kabat-Zinn, 2003). Mindfulness skills are taught in order to increase intentional attention, to develop a different relationship with one's thoughts, and to practice different strategies in relation to distressing thoughts and emotions in a non-judgmental way. MBSR (mindfulness-based stress reduction; Kabat-Zinn & Hanh, 2009) was developed as secular manualized group-based intervention program. Such manualized group interventions usually consist of eight weekly 2 to 2.5-hour classes with approximately 12 participants. In addition, 1-day retreat is often included. A key feature of	Code if the intervention includes the education in formal and informal mindfulness meditation practices to train both the attentional control as well as the non-judgemental attitudinal aspects of mindfulness. Code also if a structured group-based intervention programme was implemented, e.g. MBSR. Do not code if mindfulness interventions were implemented as part of another therapeutic approach (e.g. DBT). MBCT

	MBIs is the education in formal and informal mindfulness meditation practices to train both the attentional control as well as the non-judgemental attitudinal aspects of mindfulness. Through experiential practices and exercises, participants learn to step back or disengage from initial thoughts by creating a meta-awareness (i.e., awareness of being aware), which in turn counters the repetitive negative thinking and increases cognitive flexibility (Segal & Teasdale, 2018).	(Segal, Williams & Teasdale, 2002) is not coded here as it contains therapeutic components beyond mindfulness.
REL	Relaxation therapy is an umbrella term for techniques that helps people to be more relaxed when confronted by pain or a stressful situation. A variety of methods are used, including progressive muscle relaxation, autogenic training, mental imaging, music, and breathing retraining, to induce a natural state of relaxation. During and after relaxation, thoughts begin to flow slowly and naturally, muscle tension diminishes, and breathing slows and becomes deeper and more regular (Vickers, Zollman & Payne, 2001). This allows the parasympathetic branch of the autonomic nervous system to take over.	Code if any type of relaxation training was implemented without other therapeutic components, e.g. progressive muscle relaxation, autogenic training, mental imaging, or breathing retraining. SIT is not coded here.
PsEd	Psychoeducation (PsEd) refers to the process of providing education and information regarding PTSD to the patient and sometimes family members. The goal is to help better understand the condition, the challenges they are facing as well as the personal coping abilities and resources. An essential part of psychoeducation is explaining the patient in what ways PTSD might impact function.	Code if the client received education about health (e.g. about self-care, personal resources) and PTSD (e.g. how it may impact function). Do not code if psychoeducation was given as part of another therapeutic intervention.
WL	A treatment will be classified as a waiting list condition (WL) if the participants were repeatedly assessed without receiving any treatment or if they received delayed treatment (e.g. after the post-assessment in the experimental group).	Code if participants were repeatedly assessed without any treatment or if they received delayed treatment. Do not code if any psychological intervention was delivered or if new medication was administered. Primary care provider visits during waiting period are in line with WL classification criteria.

Note: EMDR = eye movement desensitization and reprocessing; CPT=cognitive processing therapy; CT= cognitive therapy; PE= prolonged exposure; NET= narrative exposure therapy; SIT= stress inoculation therapy; PsEd= psychoeducation; IPT= interpersonal therapy; MBI= mindfulness-based interventions; PCT= present – centered therapy; REL= relaxation; WL= waitlist

Appendix B

Search strategies

Source	Search string or reference
MEDLINE	((Stress Disorders, Post-Traumatic) AND (Psychotherapy OR Yoga OR psychoeducation.mp OR cognitive therapy.mp OR stress management.mp OR Cognitive-Behavioral Therapy OR Autogenic Training OR Biofeedback OR present-centered.mp OR interpersonal psychotherapy.mp OR cognitive processing.mp OR Relaxation OR Muscle Relaxation OR Relaxation Therapy OR Meditation OR Mindfulness OR exposure therapy.mp OR Implosive Therapy OR stress inoculation.mp OR Eye Movement Desensitization Therapy)) Limits: all adults; clinical trial, all or clinical trial protocol or clinical trial or controlled clinical trial or randomized controlled trial
PsychINFO	((Exp Posttraumatic Stress Disorder OR Exp Acute Stress Disorder OR Exp "Stress and Trauma Related Disorders" OR Exp Emotional Trauma) AND (exp Psychotherapy OR exp Yoga OR exp Psychoeducation OR exp Cognitive therapy OR exp Behavior Therapy OR exp Cognitive-Behavioral Therapy OR exp Autogenic Training OR exp Biofeedback Training or exp Biofeedback OR exp Interpersonal Psychotherapy OR present-centred.mp OR exp Relaxation or exp Progressive Relaxation Therapy OR exp Muscle Relaxation or exp Relaxation Therapy OR exp Meditation OR exp Mindfulness OR exp Exposure Therapy OR exp Stress Management OR stress inoculation therapy.mp OR exp Eye Movement Desensitization Therapy)) limit 32 to (("0300 clinical trial" or "0410 experimental replication") and "300 adulthood <age 18 yrs and older>" and yr="1990 - Current")
Pubmed	("Trauma and Stressor Related Disorders/diagnosis"[Mesh] OR "Stress Disorders, Traumatic/diagnosis"[Mesh] OR "Psychological Trauma/diagnosis"[Mesh] OR "Stress Disorders, Post-Traumatic/diagnosis"[Mesh]) AND ("Randomized Controlled Trials as Topic/methods"[Mesh] OR "Randomized Controlled Trial" [Publication Type] OR "Controlled Clinical Trial" [Publication Type] OR "Clinical Trials as Topic"[Mesh]) AND ("Psychotherapy/therapy"[Mesh] OR "Therapeutics"[Mesh] OR "therapy" [Subheading])

PSYNDEX	(posttraumatic stress disorder.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR acute stress disorder.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR PTSD.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia]) AND (psychotherapy.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR psychotherap*.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR cognitive behavioral therapy.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR cbt.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR cognitive processing.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR cognitive processing.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR cognitive therapy.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR EMDR.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR eye movement desensitization reprocessing.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR stress management.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR relaxation.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR autogenic training.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR progressive muscle relaxation.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR biofeedback.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR prolonged exposure.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR exposure therapy.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR present-centered.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR meditation.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR mindfulness.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR yoga.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR psychoeducation.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia]) AND (random*.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR control*.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR zufall*.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR treatment effectiveness.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR experimental design.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia] OR versus.mp. [mp=ti, ab, id, hw, fw, tm, cw, kp, ia])
Web of Science	(((TS=("posttraumatic stress" OR PTSD OR "acute stress disorder")) AND DOCUMENT TYPES: (Article)) AND ((TI=("cognitive-behavioral therapy" OR "cognitive therapy" OR "cognitive processing" OR "cognitive restructuring" OR "exposure therapy" OR "prolonged exposure" OR yoga OR EMDR OR "eye movement desensitization" OR Biofeedback OR Meditation OR mindfulness OR relaxation OR "interpersonal psychotherapy" OR "stress inoculation")) AND DOCUMENT TYPES: (Article)) AND ((ALL=("randomly" OR "randomized" OR "random" OR "controlled" OR "clinical Trial" OR "Control Group" OR "Control condition")) AND DOCUMENT TYPES: (Article))
U.S. National Library of Medicine (NIH)	Studies with Results | Interventional Studies | Posttraumatic Stress Disorder | "cognitive-behavioral therapy" OR "eye movement desensitization" OR "EMDR" OR "prolonged exposure" OR "meditation" OR mindfulness OR relaxation | Adult, Older Adult

WHO International Clinical Trials Registry Platform	condition	Posttraumatic stress disorder OR post-traumatic stress disorder OR acute stress disorder AND
	intervention	cognitive-behavioral therapy OR cognitive therapy OR cognitive processing OR mindfulness OR meditation OR yoga OR relaxation OR eye movement desensitization OR interpersonal psychotherapy OR stress inoculation OR narrative OR exposure OR present-centered

PTSDpubs	main subject	"posttraumatic stress disorder" OR "post-traumatic stress disorder" OR "acute stress disorder" OR "PTSD" AND
	in title	"cognitive-behavioral" OR "cognitive therapy" OR "cognitive processing" OR "cognitive restructuring" OR "eye movement desensitization" OR EMDR OR "narrative exposure therapy" OR "stress inoculation" OR relaxation OR mindfulness OR meditation OR yoga OR "interpersonal psychotherapy" OR "present-centered" AND
	all fields except Fulltext- search	randomized OR random OR controlled OR control OR effectiveness OR efficacy
	publ. date	Published after 01.01.1990
	document type	Doctoral dissertation OR journal article OR Master's thesis
OATD	Subject/keywords abstract Limits	Any of these words: posttraumatic Any of these words: PTSD doctoral
Cochrane CENTRAL		((MeSH descriptor: [Stress Disorders, Post-Traumatic] explode all trees and with qualifier(s): [diagnosis - DI]) AND (MeSH descriptor: [Psychotherapy] explode all trees and with qualifier(s): [methods - MT]) OR (MeSH descriptor: [Cognitive Behavioral Therapy] explode all trees and with qualifier(s): [methods - MT]) OR (MeSH descriptor: [Mindfulness] 1 tree(s) exploded and with qualifier(s): [methods - MT]) OR (MeSH descriptor: [Meditation] explode all trees and with qualifier(s): [methods - MT]) OR (MeSH descriptor: [Biofeedback, Psychology] 3 tree(s) exploded and with qualifier(s): [methods - MT]) OR (MeSH descriptor: [Eye Movement Desensitization Reprocessing] explode all trees) OR "exposure therapy" OR "interpersonal psychotherapy" OR "cognitive processing therapy" OR "stress inoculation" OR (MeSH descriptor: [Yoga] 3 tree(s) exploded) OR "prolonged exposure" OR "narrative exposure")
Reference lists of (network) meta-analyses		Bisson, J. I., Roberts, N. P., Andrew, M., Cooper, R., & Lewis, C. (2013). Psychological therapies for chronic post-traumatic stress disorder (PTSD) in adults. *Cochrane Database of Systematic Reviews*. https://www.doi.org/10.1002/14651858.cd003388.pub4
		Chen, L., Zhang, G., Hu, M., & Liang, X. (2015). Eye movement desensitization and reprocessing versus cognitive-behavioral therapy for adult posttraumatic stress disorder: systematic review and meta-analysis. *The Journal of nervous and mental disease, 203*(6), 443-451.
		Cramer, H., Anheyer, D., Saha, F. J., & Dobos, G. (2018). Yoga for posttraumatic stress disorder–a systematic review and meta-analysis. *BMC psychiatry, 18*(1), 72.

Cusack, K., Jonas, D. E., Forneris, C. A., Wines, C., Sonis, J., Middleton, J. C., ... & Weil, A. (2016). Psychological treatments for adults with posttraumatic stress disorder: A systematic review and meta-analysis. *Clinical psychology review, 43*, 128-141.

Furuta, M., Horsch, A., Ng, E. S., Bick, D., Spain, D., & Sin, J. (2018). Effectiveness of trauma-focused psychological therapies for treating post-traumatic stress disorder symptoms in women following childbirth: a systematic review and meta-analysis. *Frontiers in psychiatry, 9*, 591.

Gallegos, A. M., Crean, H. F., Pigeon, W. R., & Heffner, K. L. (2017). Meditation and yoga for posttraumatic stress disorder: a meta-analytic review of randomized controlled trials. *Clinical psychology review, 58*, 115-124.

Gerger, H., Munder, T., Gemperli, A., Nüesch, E., Trelle, S., Jüni, P., & Barth, J. (2014). Integrating fragmented evidence by network meta-analysis: Relative effectiveness of psychological interventions for adults with post-traumatic stress disorder. Psychological Medicine, 44(15), 3151-3164.

Hopwood, T. L., & Schutte, N. S. (2017). A meta-analytic investigation of the impact of mindfulness-based interventions on post traumatic stress. *Clinical Psychology Review, 57*, 12-20.

Lee, D. J., Schnitzlein, C. W., Wolf, J. P., Vythilingam, M., Rasmusson, A. M., & Hoge, C. W. (2016). Psychotherapy versus pharmacotherapy for posttraumatic stress disorder: Systemic review and meta-analyses to determine first-line treatments. *Depression and anxiety, 33*(9), 792-806.

Roberts, N. P., Roberts, P. A., Jones, N., & Bisson, J. I. (2015). Psychological interventions for post-traumatic stress disorder and comorbid substance use disorder: A systematic review and meta-analysis. *Clinical Psychology Review, 38*, 25–38. https://www.doi.org/10.1016/j.cpr.2015.02.007

Tran, U. S., & Gregor, B. (2016). The relative efficacy of bona fide psychotherapies for post-traumatic stress disorder: a meta-analytical evaluation of randomized controlled trials. *BMC psychiatry, 16*(1), 266.

Appendix C

Results of individual studies included in initial NMA

First author, year	TR 1	TR 2	SMD (SE)
Acarturk 2015	EMDR	WL	-1.65 (0.44)
Acarturk 2016	EMDR	WL	-2.14 (0.30)
Alghamdi 2015	NET	WL	-0.39 (0.35)
Bolton 2014	CPT	WL	-0.48 (0.16)
Bormann 2018	MBI	PCT	-0.41 (0.15)
Bränström 2010	MBI	WL	-0.49 (0.28)
Bryant 2008	PE	WL	-0.92 (0.27)
Bryant 2008	CT	WL	-0.49 (0.26)
Bryant 2008	CT	PE	-0.44 (0.26)
Carletto 2016	EMDR	REL	0.07 (0.31)
Carlson 1998	EMDR	REL	-0.46 (0.43)
Duffy 2007	CT	WL	-0.88 (0.28)
Edmond 1999	EMDR	WL	-1.07 (0.35)
Ehlers 2003	CT	WL	-1.22 (0.30)
Ehlers 2004	CT	WL	-1.40 (0.43)
Ehlers 2014	CT	WL	-1.51 (0.29)
Foa 1999	PE	WL	-1.82 (0.40)
Foa 1999	SIT	WL	-1.67 (0.41)
Foa 1999	PE	SIT	-0.14 (0.31)
Foa 2005	PE	WL	-0.66 (0.23)
Foa 2018	PE	WL	-0.73 (0.20)
Foa 2018	PCT	WL	-0.63 (0.20)
Foa 2018	PE	PCT	-0.10 (0.14)
Ford 2011	PCT	WL	-1.01 (0.22)
Galovski 2013	CPT	WL	-1.42 (0.23)
Gerbarg 2013	MBI	WL	-0.76 (0.42)
Ghafoori 2016	PsEd	WL	0.60 (0.25)
Ghafoori 2017	PCT	PE	-0.04 (0.25)
Goldstein 2017	MBI	WL	-0.45 (0.30)
Hensel-Dittmann 2011	NET	SIT	-0.25 (0.44)
Hijazi 2012	NET	WL	-0.33 (0.29)

Hijazi 2014	NET	WL	-0.26 (0.27)
Högberg 2007	EMDR	WL	-0.61 (0.45)
Ironson 2002	PE	EMDR	0.62 (0.47)
Jarero 2015	EMDR	WL	-4.90 (0.85)
Jensen 1994	EMDR	WL	-0.97 (0.43)
Jindani 2015	MBI	WL	-1.06 (0.31)
Kelly 2016	MBI	WL	-0.39 (0.32)
Kim 2013	MBI	WL	-1.37 (0.48)
Kubany 2003	CT	WL	-2.13 (0.42)
Kubany 2004	CT	WL	-1.45 (0.20)
Markowitz 2015	PE	REL	-0.32 (0.24)
Markowitz 2015	IPT	REL	-0.24 (0.24)
Markowitz 2015	PE	IPT	-0.08 (0.23)
Marks 1998	CT	REL	-0.08 (0.33)
McDonagh 2005	PCT	WL	-0.88 (0.31)
Meffert 2014	IPT	WL	-1.44 (0.55)
Mitchell 2014	MBI	WL	-0.00 (0.32)
Monson 2006	CPT	WL	-1.13 (0.28)
Morath 2014	NET	WL	-0.94 (0.34)
Neuner 2004	NET	PsEd	-0.19 (0.38)
Nidich 2018	MBI	PsEd	-0.35 (0.17)
Nidich 2018	PE	PsEd	-0.18 (0.17)
Nidich 2018	PE	MBI	0.17 (0.17)
Pacella 2012	PE	WL	-0.74 (0.26)
Pearson 2019	CPT	WL	-1.50 (0.27)
Power 2002	EMDR	WL	-1.69 (0.33)
Ratcliff 2016	MBI	WL	0.11 (0.20)
Rauch 2015	PE	PCT	-0.92 (0.42)
Reger 2016	PE	WL	-0.83 (0.24)
Reinhardt 2018	MBI	WL	-0.11 (0.53)
Resick 2002	CPT	WL	-1.03 (0.21)
Resick 2002	PE	WL	-0.83 (0.20)
Resick 2002	CPT	PE	-0.20 (0.18)
Resick 2008	CPT	CT	0.10 (0.20)
Roth 2014	REL	WL	-0.12 (0.23)

Rothbaum 1997	EMDR	WL	-2.66 (0.69)
Rothbaum 2005	PE	WL	-1.81 (0.38)
Rothbaum 2005	EMDR	WL	-1.38 (0.36)
Rothbaum 2005	PE	EMDR	-0.44 (0.32)
Rothbaum 2012	PE	WL	-0.25 (0.17)
Schnurr 2007	PE	PCT	-0.24 (0.12)
Seppälä 2014	MBI	WL	-0.54 (0.45)
Shalev 2012	PE	WL	-0.85 (0.17)
Shalev 2012	CT	WL	-0.81 (0.20)
Shalev 2012	CT	PE	-0.03 (0.20)
Shapiro 2015	EMDR	WL	-1.34 (0.55)
Shapiro 2018	EMDR	WL	-0.74 (0.42)
Suris 2013	CPT	PCT	-1.07 (0.25)
Thorp 2019	PE	REL	-0.67 (0.25)
van den Berg 2015	PE	WL	-0.75 (0.21)
van den Berg 2015	EMDR	WL	-0.65 (0.20)
van den Berg 2015	PE	EMDR	-0.10 (0.19)
van der Kolk 2014	MBI	PsEd	-0.55 (0.25)
Vaughan 1995	EMDR	WL	-1.44 (0.43)
Wahbeh 2016	MBI	REL	-0.40 (0.28)
Wells 2015	PE	WL	-1.77 (0.55)
Zang 2013	NET	WL	-1.67 (0.41)
Zang 2014	NET	WL	-4.01 (0.84)

Note: TR 1= treatment the first group received; TR 2= comparison treatment; EMDR = eye movement desensitization and reprocessing; CPT=cognitive processing therapy; CT= cognitive therapy; PE= prolonged exposure; NET= narrative exposure therapy; SIT= stress inoculation therapy; PsEd= psycho education; IPT= interpersonal therapy; MBI= mindfulness-based interventions; PCT= present – centered therapy; REL= relaxation; WL= waitlist

Appendix D

Characteristics of included studies

First author	Year	Sample PTSD severity	Sample type	Type of measure	TR1 1	TR 2	TR 3	N	Trial size	Type of analysis
Acarturk	2015	clinical	refugees	self-rated	EMDR	WL		29	very small	Compl.
Acarturk	2016	clinical	refugees	self-rated	EMDR	WL		70	small-moderate	ITT
Adenauer	2011	clinical	refugees	clinician-rated	NET	WL		19	very small	Compl.
Alghamdi	2015	clinical	civilian	self-rated	NET	WL		34	very small	Compl.
Bichescu	2007	clinical	civilian	clinician-rated	NET	PsEd		18	very small	Compl.
Bolton	2014	mixed	civilian	clinician-rated	CPT	WL		167	large	ITT
Bormann	2008	clinical	military	clinician-rated	MBI	WL		29	very small	Compl.
Bormann	2018	clinical	military	self-rated	MBI	PCT		173	large	ITT
Bränström	2010	n.a.	civilian	self-rated	MBI	WL		71	small-moderate	ITT
Bränström	2012	n.a.	civilian	self-rated	MBI	WL		60	very small	Compl.
Bryant	2008	subclinical	civilian	clinician-rated	PE	CT	WL	90	small-moderate	ITT
Carletto	2016	clinical	civilian	self-rated	EMDR	REL		42	very small	Compl.
Carlson	1998	clinical	military	self-rated	EMDR	REL		23	very small	Compl.
Cigrang	2017	mixed	military	clinician-rated	PE	WL		67	small-moderate	ITT
de Bont	2013	clinical	civilian	self-rated	PE	EMDR		10	very small	ITT
Duffy	2007	clinical	mixed	self-rated	CT	WL		58	very small	ITT

								N		
Edmond	1999	clinical	civilian	self-rated	EMDR	WL		39	very small	Compl.
Edmond	2000	n.a.	civilian	self-rated	EMDR	WL		39	very small	Compl.
Ehlers	2003	clinical	civilian	self-rated	CT	WL		55	very small	Compl.
Ehlers	2004	clinical	civilian	self-rated	CT	WL		28	very small	ITT
Ehlers	2014	clinical	civilian	self-rated	CT	WL		61	small-moderate	ITT
Foa	1999	clinical	civilian	clinician-rated	PE	SIT	WL	57	very small	Compl.
Foa	2005	clinical	civilian	clinician-rated	PE	WL		105	very small	ITT
Foa	2018	clinical	military	self-rated	PE	PCT	WL	256	large	ITT
Fonzo	2017	clinical	mixed	self-rated	PE	WL		66	small-moderate	ITT
Ford	2011	clinical	civilian	clinician-rated	PCT	WL		98	small-moderate	ITT
Galovski	2013	clinical	civilian	self-rated	CPT	WL		100	small-moderate	ITT
Gerbarg	2013	clinical	military	self-rated	MBI	WL		25	very small	ITT
Ghafoori	2016	clinical	civilian	self-rated	PsEd	WL		67	small-moderate	Compl.
Ghafoori	2017	clinical	civilian	self-rated	PE	PCT		71	very small	ITT
Goldstein	2017	mixed	military	clinician-rated	MBI	WL		47	very small	ITT
Hensel-Dittmann	2011	clinical	refugees	clinician-rated	NET	SIT		21	very small	Compl.
Hijazi	2012	clinical	refugees	self-rated	NET	WL		53	very small	ITT
Hijazi	2014	mixed	refugees	clinician-rated	NET	WL		63	very small	ITT
Högberg	2007	clinical	civilian	self-rated	EMDR	WL		21	very small	Compl.

Author	Year	Status	Population	Rating	Cond1	Cond2	Cond3	N	Effect	Analysis
Ironson	2002	clinical	civilian	self-rated	PE	EMDR		19	very small	Compl.
Jacob	2014	clinical	civilian	clinician-rated	NET	WL		76	small-moderate	ITT
Jarero	2015	subclinical	civilian	self-rated	EMDR	WL		25	very small	Compl.
Jensen	1994	clinical	military	clinician-rated	EMDR	WL		25	very small	Compl.
Jindani	2015	clinical	civilian	self-rated	MBI	WL		50	very small	Compl.
Kelly	2016	mixed	civilian	self-rated	MBI	WL		39	very small	Compl.
Kim	2013	subclinical	civilian	self-rated	MBI	WL		22	very small	Compl.
Kubany	2003	clinical	civilian	self-rated	CT	WL		32	very small	Compl.
Kubany	2004	clinical	civilian	self-rated	CT	WL		125	large	ITT
Lang	2019	clinical	military	self-rated	MBI	REL		28	very small	Compl.
Markowitz	2015	clinical	civilian	self-rated	PE	IPT	REL	110	small-moderate	ITT
Marks	1998	clinical	civilian	clinician-rated	CT	REL		37	very small	Compl.
McDonagh	2005	clinical	civilian	clinician-rated	PCT	WL		45	very small	ITT
Meffert	2014	clinical	refugees	clinician-rated	IPT	WL		18	very small	Compl.
Mitchell	2014	mixed	mixed	self-rated	MBI	WL		38	very small	ITT
Monson	2006	clinical	military	self-rated	CPT	WL		60	small-moderate	ITT
Morath	2014	clinical	refugees	clinician-rated	NET	WL		38	very small	ITT
Neuner	2004	clinical	refugees	self-rated	NET	PsEd		29	very small	ITT
Neuner	2008	clinical	refugees	self-rated	NET	WL			small-moderate	Compl.

Nidich	2018	clinical	military	self-rated	PE	MBI	PsEd	202	large	ITT
Onyut	2008	clinical	refugees	self-rated	NET	WL			small-moderate	Compl.
Pacella	2012	clinical	civilian	clinician-rated	PE	WL		65	very small	ITT
Pearson	2019	clinical	civilian	self-rated	CPT	WL		73	small-moderate	ITT
Power	2002	clinical	civilian	self-rated	EMDR	WL		51	very small	Compl.
Ratcliff	2016	n.a.	civilian	self-rated	MBI	WL		97	small-moderate	Compl.
Rauch	2015	clinical	military	clinician-rated	PE	PCT		26	very small	Compl.
Reger	2016	clinical	military	self-rated	PE	WL		78	small-moderate	Compl.
Reinhardt	2018	clinical	military	self-rated	MBI	WL		15	very small	Compl.
Resick	2002	clinical	civilian	self-rated	CPT	PE	WL	171	large	ITT
Resick	2008	clinical	civilian	self-rated	CPT	CT		100	small-moderate	ITT
Rice	2018	mixed	military	self-rated	MBI	WL		89	small-moderate	Compl.
Roth	2014	mixed	military	clinician-rated	REL	WL		80	small-moderate	ITT
Rothbaum	2012	subclinical	civilian	self-rated	PE	WL		137	large	ITT
Rothbaum	2005	clinical	civilian	self-rated	PE	EMDR	WL	40	very small	Compl.
Rothbaum	1997	clinical	civilian	self-rated	EMDR	WL		18	very small	Compl.
Schnurr	2007	clinical	military	self-rated	PE	PCT		284	large	ITT
Seppälä	2014	n.a.	military	self-rated	MBI	WL		21	very small	ITT
Shalev	2012	mixed	civilian	self-rated	PE	CT	WL	196	large	ITT
Shapiro	2015	subclinical	civilian	self-rated	EMDR	WL		17	very small	Compl.

Shapiro	2018	subclinical	civilian	self-rated	EMDR	WL		24	very small	Compl.
Suris	2013	clinical	military	self-rated	CPT	PCT		86	small-moderate	ITT
Taylor	2003	clinical	civilian	clinician-rated	EMDR	REL		38	very small	ITT
Thorp	2019	clinical	military	self-rated	PE	REL		62	small-moderate	Compl.
van den Berg	2015	clinical	civilian	self-rated	PE	EMDR	WL	155	small-moderate	ITT
van der Kolk	2014	clinical	civilian	self-rated	MBI	PsEd		64	small-moderate	ITT
Vaughan	1994	mixed	civilian	clinician-rated	EMDR	WL		29	very small	Compl.
Wahbeh	2016	clinical	military	self-rated	MBI	REL		52	very small	Compl.
Wells	2015	clinical	civilian	self-rated	PE	WL		20	very small	Compl.
Zang	2014	clinical	civilian	self-rated	NET	WL		20	very small	Compl.
Zang	2013	clinical	civilian	self-rated	NET	WL		22	very small	Compl.

Note: military = active duty military or veterans; TF= trauma-focused; WL= waiting list; SM= stress management; MBIs= mindfulness-based interventions (e.g. meditation programs); CT= cognitive therapy; CPT= cognitive processing therapy; PE= prolonged exposure; SIT= stress inoculation therapy; REL= relaxation; EMDR= eye movement desensitization and reprocessing; P-placebo= psychological placebo; IPT= interpersonal therapy; PsEd= psychoeducation; PCT= present-centered therapy; NET= narrative exposure therapy; N = number of participants; ITT = intention-to-treat analysis; Compl. = analysis included only completers

Appendix E

Reference list of studies excluded because unable to obtain full text

Daroff, L. H. (1997). *Efficacy of eye movement desensitization and reprocessing procedure in the treatment of traumatic memories: A replication study* [Doctoral dissertation, Temple University]. ProQuest Dissertations and Theses database. https://search.proquest.com/docview/42452958?accountid=14652

Opdyke, D. C. (1996). *Clinical efficacy of eye movement desensitization and reprocessing (EMDR) in the treatment of posttraumatic stress disorder* [Doctoral dissertation, Georgia State University]. ProQuest Dissertations and Theses database. https://search.proquest.com/docview/42461229?accountid=14652

Swiney, U. M. (2004). *The efficacy of EMDR for survivors of a natural disaster: Intervention after Hurricane Floyd* [Doctoral dissertation, The University of North Carolina at Chapel Hill]. ProQuest Dissertations and Theses database. https://search.proquest.com/docview/42453942?accountid=14652

Appendix F

References of studies excluded after full-text review

Bormann, J. E., Liu, L., Thorp, S. R., & Lang, A. J. (2012). Spiritual wellbeing mediates PTSD

    change in veterans with military-related PTSD. *International journal of behavioral*

    medicine, *19*(4), 496-502. https://doi.org/10.1007/s12529-011-9186-1

Boukezzi, S., El Khoury-Malhame, M., Auzias, G., Reynaud, E., Rousseau, P. F., Richard, E., ... &

    Guyon, V. (2017). Grey matter density changes of structures involved in Posttraumatic Stress

    Disorder (PTSD) after recovery following Eye Movement Desensitization and Reprocessing

    (EMDR) therapy. *Psychiatry Research: Neuroimaging*, *266*, 146-152.

    https://doi.org/10.1016/j.pscychresns.2017.06.009

Bremner, J. D., Mishra, S., Campanella, C., Shah, M., Kasher, N., Evans, S., ... & Vaccarino, V.

    (2017). A pilot study of the effects of mindfulness-based stress reduction on post-traumatic

    stress disorder symptoms and brain response to traumatic reminders of combat in operation

    enduring Freedom/Operation Iraqi freedom combat veterans with post-traumatic stress

    disorder. *Frontiers in psychiatry*, *8*, 157. https://doi.org/10.3389/fpsyt.2017.00157

Bryan, C. J., Clemans, T. A., Hernandez, A. M., Mintz, J., Peterson, A. L., Yarvis, J. S., ... &

    STRONG STAR Consortium. (2016). Evaluating potential iatrogenic suicide risk in trauma-

    focused group cognitive behavioral therapy for the treatment of PTSD in active duty military

    personnel. *Depression and anxiety*, *33*(6), 549-557. https://doi.org/10.1002/da.22456

Carlsson, J., Sonne, C., Vindbjerg, E., & Mortensen, E. L. (2018). Stress management versus

    cognitive restructuring in trauma-affected refugees—A pragmatic randomised

    study. *Psychiatry research*, *266*, 116-123. https://doi.org/10.1016/j.psychres.2018.05.015

Chard, K. M. (2005). An evaluation of cognitive processing therapy for the treatment of

    posttraumatic stress disorder related to childhood sexual abuse. *Journal of consulting and*

    clinical psychology, *73*(5), 965. https://doi.org/10.1037/0022-006X.73.5.965

Descilo, T., Vedamurtachar, A., Gerbarg, P. L., Nagaraja, D., Gangadhar, B. N., Damodaran, B., ... & Brown, R. P. (2010). Effects of a yoga breath intervention alone and in combination with an exposure therapy for post-traumatic stress disorder and depression in survivors of the 2004 South-East Asia tsunami. *Acta Psychiatrica Scandinavica*, *121*(4), 289-300. https://doi.org/10.1111/j.1600-0447.2009.01466.x

Dick, A. M., Niles, B. L., Street, A. E., DiMartino, D. M., & Mitchell, K. S. (2014). Examining mechanisms of change in a yoga intervention for women: The influence of mindfulness, psychological flexibility, and emotion regulation on PTSD symptoms. *Journal of Clinical Psychology*, *70*(12), 1170-1182. https://doi.org/10.1002/jclp.22104

Doering, S., Ohlmeier, M. C., de Jongh, A., Hofmann, A., & Bisping, V. (2013). Efficacy of a trauma-focused treatment approach for dental phobia: a randomized clinical trial. *European journal of oral sciences, 121*(6), 584-593. https://doi.org/10.1111/eos.12090

Edmond, T. E. (2000, April 14). *Eye movement desensitization and reprocessing: evaluating its effectiveness in reducing trauma symptoms in adult female survivors of childhood sexual abuse* [Conference presentation]. 12th National Symposium on Doctoral Research in Social Work. http://hdl.handle.net/1811/36904

Ertl, V., Pfeiffer, A., Schauer, E., Elbert, T., & Neuner, F. (2011). Community-implemented trauma therapy for former child soldiers in Northern Uganda: a randomized controlled trial. *Jama*, *306*(5), 503-512. https://doi.org/10.1001/jama.2011.1060

Feeny, N. C., Zoellner, L. A., & Foa, E. B. (2002). Treatment outcome for chronic PTSD among female assault victims with borderline personality characteristics: A preliminary examination. *Journal of Personality Disorders*, *16*(1), 30-40. https://doi.org/10.1521/pedi.16.1.30.22555

Foa, E. B., Rothbaum, B. O., Riggs, D. S., & Murdock, T. B. (1991). Treatment of posttraumatic stress disorder in rape victims: a comparison between cognitive-behavioral procedures and

counseling. *Journal of consulting and clinical psychology*, *59*(5), 715.

https://doi.org/10.1037/0022-006X.59.5.715

Fonzo, G. A., Goodkind, M. S., Oathes, D. J., Zaiko, Y. V., Harvey, M., Peng, K. K., ... & Arnow, B.

A. (2017). PTSD psychotherapy outcome predicted by brain activation during emotional

reactivity and regulation. *American Journal of Psychiatry*, *174*(12), 1163-1174.

https://doi.org/10.1176/appi.ajp.2017.16091072

Fonzo, G. A., Goodkind, M. S., Oathes, D. J., Zaiko, Y. V., Harvey, M., Peng, K. K., ... &

Rosenberg, B. M. (2017). Selective effects of psychotherapy on frontopolar cortical function

in PTSD. *American Journal of Psychiatry*, *174*(12), 1175-1184.

https://doi.org/10.1176/appi.ajp.2017.16091073

Galatzer-Levy, I. R., Ankri, Y., Freedman, S., Israeli-Shalev, Y., Roitman, P., Gilad, M., & Shalev,

A. Y. (2013). Early PTSD symptom trajectories: persistence, recovery, and response to

treatment: results from the Jerusalem Trauma Outreach and Prevention Study (J-TOPS). *PloS*

one, *8*(8). https://doi.org/10.1371/journal.pone.0070084

Gallagher, M. W., & Resick, P. A. (2012). Mechanisms of change in cognitive processing therapy

and prolonged exposure therapy for PTSD: Preliminary evidence for the differential effects of

hopelessness and habituation. *Cognitive therapy and research*, *36*(6), 750-755.

https://doi.org/10.1007/s10608-011-9423-6

Galovski, T. E., Monson, C., Bruce, S. E., & Resick, P. A. (2009). Does cognitive–behavioral

therapy for PTSD improve perceived health and sleep impairment?. *Journal of Traumatic*

Stress: Official Publication of The International Society for Traumatic Stress Studies, *22*(3),

197-204. https://doi.org/10.1002/jts.20418

Gradus, J. L., Suvak, M. K., Wisco, B. E., Marx, B. P., & Resick, P. A. (2013). Treatment of

posttraumatic stress disorder reduces suicidal ideation. *Depression and Anxiety*, *30*(10), 1046-

1053. https://doi.org/10.1002/da.22117

Gutner, C. A., Nillni, Y. I., Suvak, M., Wiltsey-Stirman, S., & Resick, P. A. (2013). Longitudinal

course of anxiety sensitivity and PTSD symptoms in cognitive-behavioral therapies for

PTSD. *Journal of anxiety disorders*, *27*(7), 728-734.

https://doi.org/10.1016/j.janxdis.2013.09.010

Gutner, C. A., Casement, M. D., Gilbert, K. S., & Resick, P. A. (2013). Change in sleep symptoms

across cognitive processing therapy and prolonged exposure: a longitudinal

perspective. *Behaviour research and therapy*, *51*(12), 817-822.

https://doi.org/10.1016/j.brat.2013.09.008

Gutner, C. A., Suvak, M. K., Sloan, D. M., & Resick, P. A. (2016). Does timing matter? Examining

the impact of session timing on outcome. *Journal of consulting and clinical

psychology*, *84*(12), 1108. https://doi.org/10.1037/ccp0000120

Gutner, C. A., Gallagher, M. W., Baker, A. S., Sloan, D. M., & Resick, P. A. (2016). Time course of

treatment dropout in cognitive–behavioral therapies for posttraumatic stress

disorder. *Psychological Trauma: Theory, Research, Practice, and Policy*, *8*(1), 115.

https://doi.org/10.1037/tra0000062

Harris, J. I., Usset, T., Voecks, C., Thuras, P., Currier, J., & Erbes, C. (2018). Spiritually integrated

care for PTSD: A randomized controlled trial of "Building Spiritual Strength". *Psychiatry

research*, *267*, 420-428. https://doi.org/10.1016/j.psychres.2018.06.045

Heffner, K. L., Crean, H. F., & Kemp, J. E. (2016). Meditation programs for veterans with

posttraumatic stress disorder: Aggregate findings from a multi-site evaluation. *Psychological

Trauma: Theory, Research, Practice, and Policy, 8*(3), 365–

374. https://doi.org/10.1037/tra0000106

Hensel-Dittmann, D. (2007). *Kontrollierte Therapieevaluation der Narrativen Expositionstherapie

(NET) im Vergleich zu Stress-Impfungs-Training (SIT) bei posttraumatischer

Belastungsstörung in Folge organisierter Gewalt* [Doctoral dissertation, University of

Konstanz]. Konstanzer Online-Publikationssystem. http://nbn-resolving.de/urn:nbn:de:bsz:352-opus-52629

Hien, D. A., Jiang, H., Campbell, A. N., Hu, M. C., Miele, G. M., Cohen, L. R., ... & Suarez-Morales, L. (2010). Do treatment improvements in PTSD severity affect substance use outcomes? A secondary analysis from a randomized clinical trial in NIDA's Clinical Trials Network. *American Journal of Psychiatry*, *167*(1), 95-101. https://doi.org/10.1176/appi.ajp.2009.09091261

Hinsberger, M. (2017). *Psychotherapy for offender rehabilitation in a context of ongoing community and gang violence* [Doctoral dissertation, University of Konstanz]. Konstanzer Online-Publikationssystem. http://nbn-resolving.de/urn:nbn:de:bsz:352-0-405528

Hinsberger, M., Holtzhausen, L., Sommer, J., Kaminer, D., Elbert, T., Seedat, S., ... & Weierstall, R. (2017). Feasibility and effectiveness of narrative exposure therapy and cognitive behavioral therapy in a context of ongoing violence in South Africa. *Psychological trauma: theory, research, practice, and policy*, *9*(3), 282. http://dx.doi.org/10.1037/tra0000197

Holder, N., Holliday, R., Williams, R., Mullen, K., & Surís, A. (2018). A preliminary examination of the role of psychotherapist fidelity on outcomes of cognitive processing therapy during an RCT for military sexual trauma-related PTSD. *Cognitive behaviour therapy*, *47*(1), 76-89. https://doi.org/10.1080/16506073.2017.1357750

Holliday, R., Link-Malcolm, J., Morris, E. E., & Surís, A. (2014). Effects of cognitive processing therapy on PTSD-related negative cognitions in veterans with military sexual trauma. *Military Medicine*, *179*(10), 1077-1082. https://doi.org/10.7205/MILMED-D-13-00309

Holliday, R., Williams, R., Bird, J., Mullen, K., & Surís, A. (2015). The role of cognitive processing therapy in improving psychosocial functioning, health, and quality of life in veterans with military sexual trauma-related posttraumatic stress disorder. *Psychological Services*, *12*(4), 428. https://doi.org/10.1037/ser0000058

Holliday, R., Holder, N., & Surís, A. (2018). Reductions in self-blame cognitions predict PTSD improvements with cognitive processing therapy for military sexual trauma-related PTSD. *Psychiatry Research*, *263*, 181-184. https://doi.org/10.1016/j.psychres.2018.03.007

Horesh, D., Qian, M., Freedman, S., & Shalev, A. (2017). Differential Effect of Prolonged Exposure Therapy and Cognitive Therapy on PTSD Symptom Clusters: A Randomized Controlled Trial. *Psychology and psychotherapy*, *90*(2), 235. https://www.doi.org/10.1111/papt.12103

Hourani, L., Tueller, S., Kizakevich, P., Strange, L., Lewis, G., Weimer, B., ... & Nelson, J. (2018). Effect of stress inoculation training with relaxation breathing on perceived stress and posttraumatic stress disorder in the military: A longitudinal study. *International Journal of Stress Management*, *25*(S1), 124. https://doi.org/10.1037/str0000082

Iverson, K. M., Gradus, J. L., Resick, P. A., Suvak, M. K., Smith, K. F., & Monson, C. M. (2011). Cognitive–behavioral therapy for PTSD and depression symptoms reduces risk for future intimate partner violence among interpersonal trauma survivors. *Journal of consulting and clinical psychology*, *79*(2), 193. https://doi.org/10.1037/a0022512

Iverson, K. M., King, M. W., Cunningham, K. C., & Resick, P. A. (2015). Rape survivors' trauma-related beliefs before and after Cognitive processing therapy: Associations with PTSD and depression symptoms. *Behaviour research and therapy*, *66*, 49-55. https://doi.org/10.1016/j.brat.2015.01.002

Jacob, N. (2010). *Consequences of traumatic stress in Rwandan genocide survivors: Epidemiology, psychotherapy, and dissemination* [Doctoral dissertation, University of Konstanz]. Konstanzer Online-Publikationssystem. http://nbn-resolving.de/urn:nbn:de:bsz:352-opus-106870

Johnson, D. R., & Lubin, H. (2006). The counting method: Applying the rule of parsimony to the treatment of posttraumatic stress disorder. *Traumatology*, *12*(1), 83-99. https://doi.org/10.1177/15347656060120010

King, A. P., Block, S. R., Sripada, R. K., Rauch, S., Giardino, N., Favorite, T., ... & Liberzon, I. (2016). Altered default mode network (DMN) resting state functional connectivity following

a mindfulness-based exposure therapy for posttraumatic stress disorder (PTSD) in combat veterans of Afghanistan and Iraq. *Depression and anxiety*, *33*(4), 289-299. https://doi.org/10.1002/da.22481

Korte, K. J., Allan, N. P., Gros, D. F., & Acierno, R. (2016). Differential treatment response trajectories in individuals with subclinical and clinical PTSD. *Journal of anxiety disorders*, *38*, 95-101. https://doi.org/10.1016/j.janxdis.2016.01.006

Lang, A. J., Malaktaris, A. L., Casmar, P., Baca, S. A., Golshan, S., Harrison, T., & Negi, L. (2019). Compassion meditation for posttraumatic stress disorder in veterans: A randomized proof of concept study. *Journal of traumatic stress*, *32*(2), 299-309. https://doi.org/10.1002/jts.22397

Larsen, S. E., Fleming, C. J., & Resick, P. A. (2019). Residual symptoms following empirically supported treatment for PTSD. *Psychological trauma: theory, research, practice, and policy*, *11*(2), 207. https://doi.org/10.1037/tra0000384

Lee, M. Y., Zaharlick, A., & Akers, D. (2017). Impact of meditation on mental health outcomes of female trauma survivors of interpersonal violence with co-occurring disorders: A randomized controlled trial. *Journal of interpersonal violence*, *32*(14), 2139-2165. https://doi.org/10.1177/0886260515591277

Leiner, A. S., Kearns, M. C., Jackson, J. L., Astin, M. C., & Rothbaum, B. O. (2012). Avoidant coping and treatment outcome in rape-related posttraumatic stress disorder. *Journal of consulting and clinical psychology*, *80*(2), 317. https://doi.org/10.1037/a0026814

Lenferink, L. I., Piersma, E., de Keijser, J., Smid, G. E., & Boelen, P. A. (2017). Cognitive therapy and eye movement desensitization and reprocessing for reducing psychopathology among disaster-bereaved individuals: study protocol for a randomized controlled trial. *European journal of psychotraumatology*, *8*(1), 1388710. https://doi.org/10.1080/20008198.2017.1388710

Lewis, G. F., Hourani, L., Tueller, S., Kizakevich, P., Bryant, S., Weimer, B., & Strange, L. (2015). Relaxation training assisted by heart rate variability biofeedback: Implication for a military

predeployment stress inoculation protocol. *Psychophysiology*, *52*(9), 1167-1174. https://doi.org/10.1111/psyp.12455

Livanou, M., BaoŞĞLu, M., Marks, I. M., De Silva, P., Noshirvani, H., Lovell, K., & Thrasher, S. (2002). Beliefs, sense of control and treatment outcome in post-traumatic stress disorder. *Psychological Medicine*, *32*(1), 157-165. https://doi.org/10.1017/S0033291701004767

Liverant, G. I., Suvak, M. K., Pineles, S. L., & Resick, P. A. (2012). Changes in posttraumatic stress disorder and depressive symptoms during cognitive processing therapy: Evidence for concurrent change. *Journal of Consulting and Clinical Psychology*, *80*(6), 957. https://doi.org/10.1037/a0030485

Macdonald, A., Monson, C. M., Doron-Lamarca, S., Resick, P. A., & Palfai, T. P. (2011). Identifying patterns of symptom change during a randomized controlled trial of cognitive processing therapy for military-related posttraumatic stress disorder. *Journal of traumatic stress*, *24*(3), 268-276. https://doi.org/10.1002/jts.20642

Markowitz, J. C., Petkova, E., Biyanova, T., Ding, K., Suh, E. J., & Neria, Y. (2015). Exploring personality diagnosis stability following acute psychotherapy for chronic posttraumatic stress disorder. *Depression and anxiety*, *32*(12), 919-926. https://doi.org/10.1002/da.22436

Markowitz, J. C., Meehan, K. B., Petkova, E., Zhao, Y., Van, P. M., Neria, Y., ... & Nazia, Y. (2016). Treatment preferences of psychotherapy patients with chronic PTSD. *The Journal of clinical psychiatry*, *77*(3), 363-370. https://www.doi.org/10.4088/jcp.14m09640

Markowitz, J. C., Neria, Y., Lovell, K., Van Meter, P. E., & Petkova, E. (2017). History of sexual trauma moderates psychotherapy outcome for posttraumatic stress disorder. *Depression and anxiety*, *34*(8), 692-700. https://doi.org/10.1002/da.22619

Miller, K. E., Micol, R. L., Davis, J. L., Cranston, C. C., & Pruiksma, K. E. (2019). Predictors

of treatment noninitiation, dropout, and response for cognitive behavioral therapy for trauma nightmares. *Psychological Trauma: Theory, Research, Practice, and Policy*, *11*(1), 122. https://doi.org/10.1037/tra0000389

Monson, C. M., Gradus, J. L., Young-Xu, Y., Schnurr, P. P., Price, J. L., & Schumm, J. A. (2008). Change in posttraumatic stress disorder symptoms: do clinicians and patients agree?. *Psychological assessment*, *20*(2), 131. https://doi.org/10.1037/1040-3590.20.2.131

Monson, C. M., Macdonald, A., Vorstenbosch, V., Shnaider, P., Goldstein, E. S., Ferrier-Auerbach, A. G., & Mocciola, K. E. (2012). Changes in social adjustment with cognitive processing therapy: Effects of treatment and association with PTSD symptom change. *Journal of traumatic stress*, *25*(5), 519-526. https://doi.org/10.1002/jts.21735

Morath, J., Gola, H., Sommershof, A., Hamuni, G., Kolassa, S., Catani, C., ... & Groettrup, M. (2014). The effect of trauma-focused therapy on the altered T cell distribution in individuals with PTSD: evidence from a randomized controlled trial. *Journal of Psychiatric Research*, *54*, 1-10. https://doi.org/10.1016/j.jpsychires.2014.03.016

Mueller, B. A. (2013). *The effects of Bikram-style hot yoga on posttraumatic stress and depressive symptoms in a military population* [Doctoral Alliant International University]. ProQuest Dissertations and Theses database. https://search.proquest.com/docview/1448203661?accountid=14652

Mullen, K., Holliday, R., Morris, E., Raja, A., & Surís, A. (2014). Cognitive processing therapy for male veterans with military sexual trauma-related posttraumatic stress disorder. *Journal of anxiety disorders*, *28*(8), 761-764. https://doi.org/10.1016/j.janxdis.2014.09.004

Neuner, F. (2003). *Epidemiology and treatment of posttraumatic stress disorder in west-nile populations of sudan and uganda* [Doctoral dissertation, University of Konstanz]. Konstanzer Online-Publikationssystem. http://nbn-resolving.de/urn:nbn:de:bsz:352-opus-10828

Norr, A. M., Smolenski, D. J., & Reger, G. M. (2018). Effects of prolonged exposure and virtual reality exposure on suicidal ideation in active duty soldiers: an examination of potential

mechanisms. *Journal of psychiatric research*, *103*, 69-74.

https://doi.org/10.1016/j.jpsychires.2018.05.009

Pasternak, J. (2009). Skuteczność terapii odwrażliwiania za pomocą ruchów gałek ocznych (EMDR) u leczonych ambulatoryjnie pacjentów z zaburzeniem stresowym pourazowym. *Postêpy Psychiatrii i Neurologii, 18*(3), 247-254.

Rauch, S. A., Grunfeld, T. E., Yadin, E., Cahill, S. P., Hembree, E., & Foa, E. B. (2009). Changes in reported physical health symptoms and social function with prolonged exposure therapy for chronic posttraumatic stress disorder. *Depression and anxiety*, *26*(8), 732-738. https://doi.org/10.1002/da.20518

Reddy, S., Dick, A. M., Gerber, M. R., & Mitchell, K. (2014). The effect of a yoga intervention on alcohol and drug abuse risk in veteran and civilian women with posttraumatic stress disorder. *The Journal of Alternative and Complementary Medicine*, *20*(10), 750-756. https://doi.org/10.1089/acm.2014.0014

Renner, W., Bänninger-Huber, E., & Peltzer, K. (2011). Culture-Sensitive and Resource Oriented Peer (CROP)-Groups as a community based intervention for trauma survivors: A randomized controlled pilot study with refugees and asylum seekers from Chechnya. *Australasian Journal of Disaster and Trauma Studies*, *2011*(1), 1–13. http://hdl.handle.net/20.500.11910/3725

Resick, P. A., Nishith, P., & Griffin, M. G. (2003). How well does cognitive-behavioral therapy treat symptoms of complex PTSD? An examination of child sexual abuse survivors within a clinical trial. *CNS spectrums*, *8*(5), 340-355. https://doi.org/10.1017/S1092852900018605

Resick, P. A., Suvak, M. K., Johnides, B. D., Mitchell, K. S., & Iverson, K. M. (2012). The impact of dissociation on PTSD treatment with cognitive processing therapy. *Depression and Anxiety*, *29*(8), 718-730. https://doi.org/10.1002/da.21938

Resick, P. A., Williams, L. F., Suvak, M. K., Monson, C. M., & Gradus, J. L. (2012). Long-term outcomes of cognitive–behavioral treatments for posttraumatic stress disorder among female

rape survivors. *Journal of consulting and clinical psychology*, *80*(2), 201.

https://doi.org/10.1037/a0026602

Rice, V. J., Liu, B., & Schroeder, P. J. (2018). Impact of in-person and virtual world mindfulness

training on symptoms of post-traumatic stress disorder and attention deficit and hyperactivity

disorder. *Military medicine*, *183*(suppl_1), 413-420. https://doi.org/10.1093/milmed/usx227

Rizvi, S. L., Vogt, D. S., & Resick, P. A. (2009). Cognitive and affective predictors of treatment

outcome in cognitive processing therapy and prolonged exposure for posttraumatic stress

disorder. *Behaviour research and therapy*, *47*(9), 737-743.

https://doi.org/10.1016/j.brat.2009.06.003

Rutledge, T., Nidich, S., Schneider, R. H., Mills, P. J., Salerno, J., Heppner, P., ... & Rainforth, M.

(2014). Design and rationale of a comparative effectiveness trial evaluating transcendental

meditation against established therapies for PTSD. *Contemporary clinical trials*, *39*(1), 50-56.

https://doi.org/10.1016/j.cct.2014.07.005

Scher, C. D., Suvak, M. K., & Resick, P. A. (2017). Trauma cognitions are related to symptoms up to

10 years after cognitive behavioral treatment for posttraumatic stress disorder. *Psychological*

trauma: theory, research, practice, and policy, *9*(6), 750. https://doi.org/10.1037/tra0000258

Schnurr, P. P., Lunney, C. A., Forshay, E., Thurston, V. L., Chow, B. K., Resick, P. A., & Foa, E. B.

(2009). Sexual function outcomes in women treated for posttraumatic stress disorder. *Journal*

of women's health, *18*(10), 1549-1557. https://doi.org/10.1089/jwh.2008.1165

Schnurr, P. P., & Lunney, C. A. (2012). Work-related outcomes among female veterans and service

members after treatment of posttraumatic stress disorder. *Psychiatric Services*, *63*(11), 1072-

1079. https://doi.org/10.1176/appi.ps.201100415

Schnurr, P. P., & Lunney, C. A. (2015). Differential effects of prolonged exposure on posttraumatic

stress disorder symptoms in female veterans. *Journal of consulting and clinical*

psychology, *83*(6), 1154. https://doi.org/10.1037/ccp0000031

Schnurr, P. P., & Lunney, C. A. (2016). Symptom benchmarks of improved quality of life in PTSD. *Depression and anxiety*, *33*(3), 247-255. https://doi.org/10.1002/da.22477

Schnurr, P. P., & Lunney, C. A. (2019). Residual symptoms following prolonged exposure and present-centered therapy for PTSD in female veterans and soldiers. *Depression and anxiety*, *36*(2), 162-169. https://doi.org/10.1002/da.22871

Shalev, A. Y., Ankri, Y., Gilad, M., Israeli-Shalev, Y., Adessky, R., Qian, M., & Freedman, S. (2016). Long-term outcome of early interventions to prevent posttraumatic stress disorder. *The Journal of clinical psychiatry*, *77*(5), e580-7. https://www.doi.org/10.4088/jcp.15m09932

Sommer, J. (2017). *Social acknowledgment and its relation to PTSD and aggression in a victim-offender sample in South Africa: new perspectives on a major public health problem* [Doctoral dissertation, University of Konstanz]. Konstanzer Online-Publikationssystem. http://nbn-resolving.de/urn:nbn:de:bsz:352-0-399851

Sripada, R. K., Rauch, S. A., Tuerk, P. W., Smith, E., Defever, A. M., Mayer, R. A., ... & Venners, M. (2013). Mild traumatic brain injury and treatment response in prolonged exposure for PTSD. *Journal of Traumatic Stress*, *26*(3), 369-375. https://doi.org/10.1002/jts.21813

Stapleton, J. A., Taylor, S., & Asmundson, G. J. (2006). Effects of three PTSD treatments on anger and guilt: exposure therapy, eye movement desensitization and reprocessing, and relaxation training. *Journal of Traumatic Stress*, *19*(1), 19-28. https://doi.org/10.1002/jts.20095

Tarquinio, C., Brennstuhl, M. J., Rydberg, J. A., Schmitt, A., Mouda, F., Lourel, M., & Tarquinio, P. (2012). Eye movement desensitization and reprocessing (EMDR) therapy in the treatment of victims of domestic violence: A pilot study. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology*, *62*(4), 205-212. http://dx.doi.org/10.1016/j.erap.2012.08.006

Telles, S., Singh, N., Joshi, M., & Balkrishna, A. (2010). Post traumatic stress symptoms and heart rate variability in Bihar flood survivors following yoga: a randomized controlled study. *BMC psychiatry*, *10*(1), 18. https://doi.org/10.1186/1471-244X-10-18

Thrasher, S., Power, M., Morant, N., Marks, I., & Dalgleish, T. (2010). Social support moderates outcome in a randomized controlled trial of exposure therapy and (or) cognitive restructuring for chronic posttraumatic stress disorder. *The Canadian Journal of Psychiatry*, *55*(3), 187-190. https://doi.org/10.1177/070674371005500311

Tutarel-Kislak, S. (2004). Eye Movement Desensitization and Reprocessing (EMDR) approach in the treatment of anxiety Sensitivity. *Turk Psikoloji Dergisi*, *19*(53), 49-67.

van den Berg, D. P., de Bont, P. A., van der Vleugel, B. M., De Roos, C., de Jongh, A., van Minnen, A., & van der Gaag, M. (2016). Trauma-focused treatment in PTSD patients with psychosis: symptom exacerbation, adverse events, and revictimization. *Schizophrenia bulletin*, *42*(3), 693-702. https://doi.org/10.1093/schbul/sbv172

Van Emmerik, A. A., Kamphuis, J. H., & Emmelkamp, P. M. (2008). Treating acute stress disorder and posttraumatic stress disorder with cognitive behavioral therapy or structured writing therapy: A randomized controlled trial. *Psychotherapy and psychosomatics*, *77*(2), 93-100. https://doi.org/10.1159/000112886

Wachen, J. S., Jimenez, S., Smith, K., & Resick, P. A. (2014). Long-term functional outcomes of women receiving cognitive processing therapy and prolonged exposure. *Psychological Trauma: Theory, Research, Practice, and Policy*, *6*(S1), S58. https://doi.org/10.1037/a0035741

Walsh, K., Gilmore, A. K., Frazier, P., Ledray, L., Acierno, R., Ruggiero, K. J., ... & Resnick, H. S. (2017). A randomized clinical trial examining the effect of video-based prevention of alcohol and marijuana use among recent sexual assault victims. *Alcoholism: clinical and experimental research*, *41*(12), 2163-2172. https://doi.org/10.1111/acer.13505

Wilson, S. A., Becker, L. A., & Tinker, R. H. (1995). Eye movement desensitization and

    reprocessing (EMDR) treatment for psychologically traumatized individuals. *Journal of*

    consulting and clinical psychology, *63*(6), 928. https://doi.org/10.1037/0022-006X.63.6.928

Wilson, S. A., Tinker, R. H., Becker, L. A., & Logan, C. R. (2001). Stress management with law

    enforcement personnel: A controlled outcome study of EMDR versus a traditional stress

    management program. *International Journal of Stress Management*, *8*(3), 179-200.

    https://doi.org/10.1023/A:1011366408693

Wolf, E. J., Lunney, C. A., & Schnurr, P. P. (2016). The influence of the dissociative subtype of

    posttraumatic stress disorder on treatment efficacy in female veterans and active duty service

    members. *Journal of Consulting and Clinical Psychology*, *84*(1), 95.

    https://doi.org/10.1037/ccp0000036

Woodward, E., Hackmann, A., Wild, J., Grey, N., Clark, D. M., & Ehlers, A. (2017). Effects of

    psychotherapies for posttraumatic stress disorder on sleep disturbances: Results from a

    randomized clinical trial. *Behaviour research and therapy*, *97*, 75-85.

    https://doi.org/10.1016/j.brat.2017.07.001

Appendix G

Studies excluded from quantitative synthesis: unable to obtain relevant data

Bormann, J. E., Thorp, S., Wetherell, J. L., & Golshan, S. (2008). A spiritually based group

   intervention for combat veterans with posttraumatic stress disorder: feasibility study. *Journal*

   of Holistic Nursing, *26*(2), 109-116. https://doi.org/10.1177/0898010107311276

Cigrang, J. A., Rauch, S. A., Mintz, J., Brundige, A. R., Mitchell, J. A., Najera, E., ... & Goodie, J. L.

   (2017). Moving effective treatment for posttraumatic stress disorder to primary care: A

   randomized controlled trial with active duty military. *Families, Systems, & Health*, *35*(4),

   450. https://doi.org/10.1037/fsh0000315

de Bont, P. A., van Minnen, A., & de Jongh, A. (2013). Treating PTSD in patients with psychosis: a

   within-group controlled feasibility study examining the efficacy and safety of evidence-based

   PE and EMDR protocols. *Behavior therapy*, *44*(4), 717-730.

   https://doi.org/10.1016/j.beth.2013.07.002

Taylor, S., Thordarson, D. S., Maxfield, L., Fedoroff, I. C., Lovell, K., & Ogrodniczuk, J. (2003).

   Comparative efficacy, speed, and adverse effects of three PTSD treatments: exposure therapy,

   EMDR, and relaxation training. *Journal of consulting and clinical psychology*, *71*(2), 330.

   https://doi.org/10.1037/0022-006X.71.2.330

Appendix H1

Estimated effects of NMA including studies that report both self- and clinician-rated PTSD: results of self-rated PTSD measures

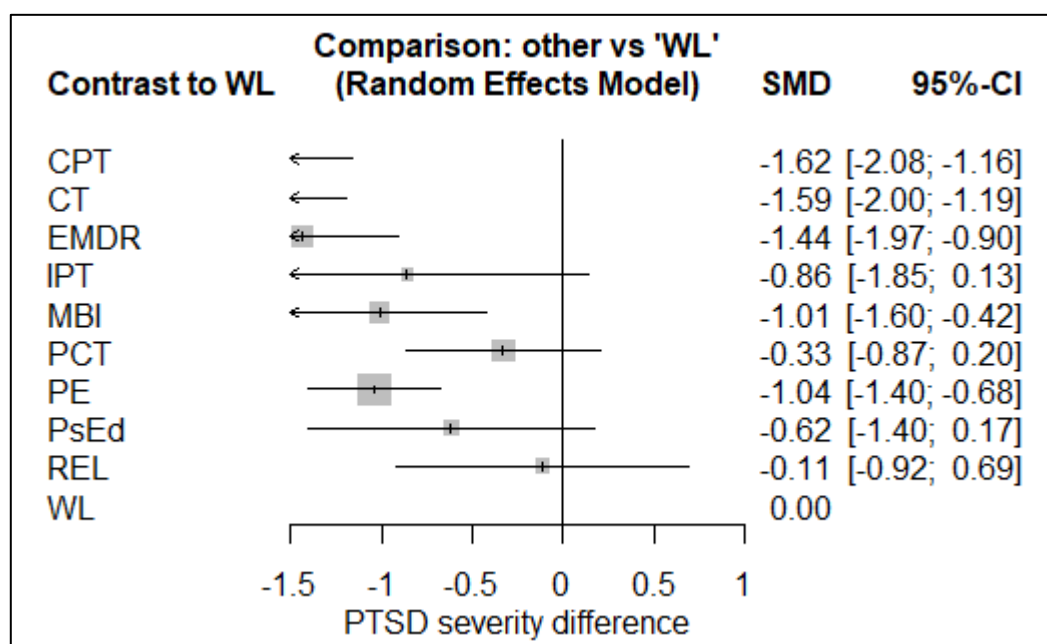| CPT									
-0.03 [-0.57; 0.52]	CT								
-0.18 [-0.87; 0.50]	-0.16 [-0.82; 0.50]	EMDR							
-0.76 [-1.82; 0.30]	-0.73 [-1.79; 0.32]	-0.58 [-1.66; 0.51]	IPT						
-0.61 [-1.31; 0.09]	-0.58 [-1.28; 0.11]	-0.43 [-1.18; 0.33]	0.15 [-0.94; 1.24]	MBI					
-1.29 [-1.90; -0.67]	-1.26 [-1.91; -0.62]	-1.10 [-1.82; -0.39]	-0.53 [-1.59; 0.53]	-0.68 [-1.30; -0.06]	PCT				
-0.58 [-1.10; -0.06]	-0.55 [-1.06; -0.05]	-0.40 [-0.96; 0.17]	0.18 [-0.75; 1.11]	0.03 [-0.55; 0.60]	0.71 [0.20; 1.22]	PE			
-1.00 [-1.87; -0.13]	-0.98 [-1.84; -0.11]	-0.82 [-1.73; 0.09]	-0.24 [-1.44; 0.95]	-0.39 [-1.06; 0.27]	0.29 [-0.54; 1.11]	-0.42 [-1.18; 0.33]	PsEd		
-1.50 [-2.39; -0.61]	-1.48 [-2.36; -0.60]	-1.32 [-2.24; -0.40]	-0.75 [-1.68; 0.19]	-0.90 [-1.82; 0.03]	-0.22 [-1.10; 0.67]	-0.93 [-1.65; -0.20]	-0.50 [-1.55; 0.54]	REL	
-1.62 [-2.08; -1.16]	-1.59 [-2.00; -1.19]	-1.44 [-1.97; -0.90]	-0.86 [-1.85; 0.13]	-1.01 [-1.60; -0.42]	-0.33 [-0.87; 0.20]	-1.04 [-1.40; -0.68]	-0.62 [-1.40; 0.17]	-0.11 [-0.92; 0.69]	WL

Note: EMDR = eye movement desensitization and reprocessing; CPT=cognitive processing therapy; CT= cognitive therapy; PE= prolonged exposure;

NET= narrative exposure therapy; SIT= stress inoculation therapy; PsEd= psychoeducation; IPT= interpersonal therapy; MBI= mindfulness-based

interventions; PCT= present – centered therapy; REL= relaxation; WL= waitlist

Appendix H2

Forest plot displaying estimated effects in the NMA of self-rated PTSD outcome measures

Contrast to WL	Comparison: other vs 'WL' (Random Effects Model)	SMD	95%-CI
CPT		-1.62	[-2.08; -1.16]
CT		-1.59	[-2.00; -1.19]
EMDR		-1.44	[-1.97; -0.90]
IPT		-0.86	[-1.85; 0.13]
MBI		-1.01	[-1.60; -0.42]
PCT		-0.33	[-0.87; 0.20]
PE		-1.04	[-1.40; -0.68]
PsEd		-0.62	[-1.40; 0.17]
REL		-0.11	[-0.92; 0.69]
WL		0.00	

PTSD severity difference

Appendix H3

League table of the NMA including studies that report both self- and clinician-rated PTSD: results of clinician-rated PTSD measures

CPT									
0.02 [-0.21; 0.26]	CT								
-0.06 [-0.56; 0.44]	-0.13 [-0.62; 0.36]	EMDR							
-0.50 [-1.27; 0.26]	-0.57 [-1.34; 0.19]	-0.44 [-1.23; 0.35]	IPT						
-0.36 [-0.86; 0.15]	-0.43 [-0.94; 0.08]	-0.30 [-0.86; 0.26]	0.14 [-0.65; 0.93]	MBI					
-0.70 [-1.12; -0.28]	-0.77 [-1.22; -0.31]	-0.64 [-1.15; -0.12]	-0.20 [-0.95; 0.56]	-0.34 [-0.78; 0.10]	PCT				
-0.34 [-0.70; 0.03]	-0.41 [-0.77; -0.04]	-0.28 [-0.69; 0.14]	0.17 [-0.51; 0.84]	0.02 [-0.39; 0.44]	0.36 [0.02; 0.71]	PE			
-0.72 [-1.34; -0.10]	-0.79 [-1.41; -0.17]	-0.66 [-1.32; 0.00]	-0.22 [-1.08; 0.64]	-0.36 [-0.83; 0.11]	-0.02 [-0.60; 0.56]	-0.38 [-0.92; 0.15]	PsEd		
-0.83 [-1.47; -0.19]	-0.90 [-1.54; -0.26]	-0.77 [-1.44; -0.10]	-0.33 [-1.01; 0.35]	-0.47 [-1.14; 0.20]	-0.13 [-0.76; 0.50]	-0.49 [-1.02; 0.03]	-0.11 [-0.86; 0.64]	REL	
-1.30 [-1.63; -0.97]	-1.37 [-1.67; -1.07]	-1.24 [-1.64; -0.84]	-0.80 [-1.52; -0.08]	-0.94 [-1.37; -0.51]	-0.60 [-0.97; -0.23]	-0.96 [-1.22; -0.70]	-0.58 [-1.14; -0.02]	-0.47 [-1.06; 0.11]	WL

Note: EMDR = eye movement desensitization and reprocessing; CPT=cognitive processing therapy; CT= cognitive therapy; PE= prolonged exposure; NET= narrative exposure therapy; SIT= stress inoculation therapy; PsEd= psychoeducation; IPT= interpersonal therapy; MBI= mindfulness-based interventions; PCT= present – centered therapy; REL= relaxation; WL= waitlist

Appendix H4

League table of NMA including all studies: results of self-rated PTSD outcome measures given precedence over clinician-rated outcome measures

CPT	CT	EMDR	IPT	MBI	NET	PCT	PE	PsEd	REL	SIT	WL
CPT											
-0.13 [-0.62; 0.37]	**CT**										
-0.06 [-0.57; 0.44]	0.06 [-0.40; 0.53]	**EMDR**									
-0.24 [-1.13; 0.64]	-0.12 [-0.98; 0.74]	-0.18 [-1.02; 0.66]	**IPT**								
-0.73 [-1.24; -0.23]	-0.61 [-1.08; -0.14]	-0.67 [-1.11; -0.23]	-0.49 [-1.34; 0.36]	**MBI**							
-0.52 [-1.13; 0.08]	-0.40 [-0.97; 0.17]	-0.46 [-1.01; 0.09]	-0.28 [-1.19; 0.64]	0.21 [-0.33; 0.75]	**NET**						
-0.96 [-1.48; -0.43]	-0.83 [-1.35; -0.32]	-0.90 [-1.39; -0.40]	-0.71 [-1.59; 0.16]	-0.22 [-0.69; 0.25]	-0.44 [-1.03; 0.16]	**PCT**					
-0.40 [-0.85; 0.05]	-0.27 [-0.68; 0.13]	-0.34 [-0.71; 0.04]	-0.15 [-0.95; 0.65]	0.34 [-0.05; 0.72]	0.12 [-0.39; 0.64]	0.56 [0.16; 0.96]	**PE**				
-1.20 [-1.88; -0.53]	-1.08 [-1.72; -0.43]	-1.14 [-1.77; -0.51]	-0.96 [-1.92; 0.00]	-0.47 [-1.03; 0.09]	-0.68 [-1.33; -0.03]	-0.24 [-0.90; 0.41]	-0.80 [-1.38; -0.23]	**PsEd**			
-0.88 [-1.46; -0.29]	-0.75 [-1.28; -0.22]	-0.81 [-1.30; -0.32]	-0.63 [-1.46; 0.20]	-0.14 [-0.65; 0.37]	-0.35 [-0.98; 0.28]	0.08 [-0.49; 0.65]	-0.48 [-0.94; -0.02]	0.33 [-0.36; 1.01]	**REL**		
-0.41 [-1.32; 0.50]	-0.29 [-1.18; 0.60]	-0.35 [-1.23; 0.53]	-0.17 [-1.31; 0.97]	0.32 [-0.55; 1.20]	0.11 [-0.74; 0.96]	0.54 [-0.36; 1.44]	-0.02 [-0.85; 0.82]	0.79 [-0.18; 1.76]	0.46 [-0.46; 1.39]	**SIT**	
-1.40 [-1.80; -1.00]	-1.27 [-1.63; -0.92]	-1.34 [-1.65; -1.02]	-1.15 [-1.95; -0.35]	-0.66 [-0.98; -0.35]	-0.88 [-1.33; -0.42]	-0.44 [-0.84; -0.04]	-1.00 [-1.26; -0.74]	-0.20 [-0.75; 0.35]	-0.52 [-0.97; -0.08]	-0.99 [-1.81; -0.16]	**WL**

Note: EMDR = eye movement desensitization and reprocessing; CPT=cognitive processing therapy; CT= cognitive therapy; PE= prolonged exposure;

NET= narrative exposure therapy; SIT= stress inoculation therapy; PsEd= psychoeducation; IPT= interpersonal therapy; MBI= mindfulness-based

interventions; PCT= present – centered therapy; REL= relaxation; WL= waitlist

Appendix I1

League table of estimated effects of NMA with results of PP analysis given precedence over those of ITT analysis, all studies included

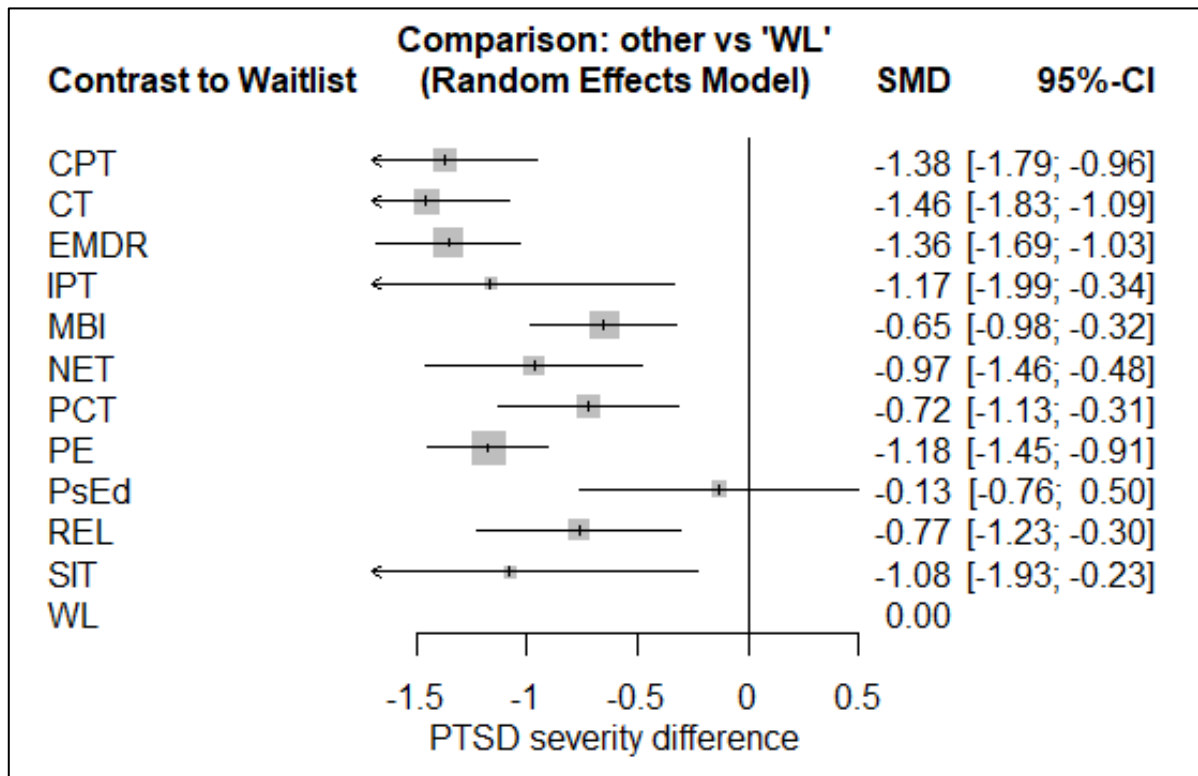| CPT											
0.08 [-0.44; 0.60]	CT										
-0.02 [-0.54; 0.51]	-0.10 [-0.58; 0.38]	EMDR									
-0.21 [-1.13; 0.71]	-0.29 [-1.18; 0.60]	-0.19 [-1.06; 0.68]	IPT								
-0.72 [-1.25; -0.20]	-0.81 [-1.30; -0.32]	-0.71 [-1.16; -0.25]	-0.51 [-1.39; 0.36]	MBI							
-0.41 [-1.05; 0.24]	-0.49 [-1.10; 0.13]	-0.39 [-0.98; 0.20]	-0.19 [-1.16; 0.77]	0.32 [-0.27; 0.91]	NET						
-0.65 [-1.19; -0.11]	-0.73 [-1.27; -0.20]	-0.64 [-1.14; -0.13]	-0.44 [-1.34; 0.46]	0.07 [-0.41; 0.56]	-0.25 [-0.89; 0.39]	PCT					
-0.20 [-0.67; 0.27]	-0.28 [-0.71; 0.15]	-0.18 [-0.57; 0.20]	0.01 [-0.82; 0.84]	0.53 [0.13; 0.92]	0.21 [-0.35; 0.76]	0.45 [0.04; 0.87]	PE				
-1.24 [-1.99; -0.50]	-1.33 [-2.05; -0.60]	-1.23 [-1.93; -0.53]	-1.03 [-2.06; -0.01]	-0.52 [-1.14; 0.11]	-0.84 [-1.64; -0.04]	-0.59 [-1.32; 0.14]	-1.04 [-1.70; -0.39]	PsEd			
-0.61 [-1.22; 0.00]	-0.69 [-1.24; -0.14]	-0.59 [-1.10; -0.09]	-0.40 [-1.25; 0.45]	0.11 [-0.41; 0.64]	-0.21 [-0.88; 0.47]	0.04 [-0.55; 0.63]	-0.41 [-0.89; 0.07]	0.63 [-0.12; 1.39]	REL		
-0.30 [-1.24; 0.65]	-0.38 [-1.30; 0.54]	-0.28 [-1.18; 0.62]	-0.09 [-1.26; 1.09]	0.43 [-0.48; 1.33]	0.11 [-0.77; 0.99]	0.36 [-0.57; 1.29]	-0.10 [-0.96; 0.76]	0.95 [-0.10; 2.00]	0.31 [-0.64; 1.27]	SIT	
-1.38 [-1.79; -0.96]	-1.46 [-1.83; -1.09]	-1.36 [-1.69; -1.03]	-1.17 [-1.99; -0.34]	-0.65 [-0.98; -0.32]	-0.97 [-1.46; -0.48]	-0.72 [-1.13; -0.31]	-1.18 [-1.45; -0.91]	-0.13 [-0.76; 0.50]	-0.77 [-1.23; -0.30]	-1.08 [-1.93; -0.23]	WL

EMDR = eye movement desensitization and reprocessing; CPT=cognitive processing therapy; CT= cognitive therapy; PE= prolonged exposure;

NET= narrative exposure therapy; SIT= stress inoculation therapy; PsEd= psychoeducation; IPT= interpersonal therapy; MBI= mindfulness-based

interventions; PCT= present – centered therapy; REL= relaxation; WL= waitlist

Appendix I2

Forest plot: NMA of all studies, results of PP analyses given precedence over ITT data

Contrast to Waitlist	Comparison: other vs 'WL' (Random Effects Model)	SMD	95%-CI
CPT		-1.38	[-1.79; -0.96]
CT		-1.46	[-1.83; -1.09]
EMDR		-1.36	[-1.69; -1.03]
IPT		-1.17	[-1.99; -0.34]
MBI		-0.65	[-0.98; -0.32]
NET		-0.97	[-1.46; -0.48]
PCT		-0.72	[-1.13; -0.31]
PE		-1.18	[-1.45; -0.91]
PsEd		-0.13	[-0.76; 0.50]
REL		-0.77	[-1.23; -0.30]
SIT		-1.08	[-1.93; -0.23]
WL		0.00	

PTSD severity difference

Appendix J

Tests for extreme effects at the level of pairwise meta-analyses: Results of studies meeting

criteria for outliers

study	rstud	dffits	cooksD	covr	τ^2 del	Qdel	hat	weight	dfbeta	inf
Jarero et al., 2015	-3.82	-1.01	0.72	0.53	0.00	10.07	0.05	4.93	-1.18	*
Zang et al., 2014	-3.54	-1.45	0.91	0.26	0.00	3.88	0.13	12.88	-1.94	*

Note: rstud = studentized deleted residuals, dffits = difference in fits, cooksD = Cook's

distance, covr = covariance ratio, τ^2 del = change in tau-squared when deleting the study, Qdel

= residual heterogeneity after deleting the study, hat = hat values, inf = * indicating that the

study is influential.

Appendix K

Heterogeneity and inconsistency in the network after outlier exclusion

	tests of heterogeneity	Q	df	p-value
Outlier exclusion at the level of pairwise meta-analyses	total	189.98	67	<0.0001
	within designs	92.78	42	<0.0001
	between designs	97.22	25	<0.0001
Network-level outlier exclusion	total	197.04	61	<0.0001
	within designs	118.87	42	<0.0001
	between designs	78.17	19	<0.0001
all outliers excluded	total	161.59	59	<0.0001
	within designs	86.62	40	<0.0001
	between designs	74.97	19	<0.0001

Note: between-design heterogeneity $\triangleq$ inconsistency

Appendix L1

Changes in treatment rankings after the exclusion of outliers

outlier exclusion at the level of pairwise MA		network-level outlier exclusion		all outliers excluded	
intervention	p-score	intervention	p-score	intervention	p-score
EMDR	0.87	EMDR	0.91	EMDR	0.88
CPT	0.85	IPT	0.81	IPT	0.84
CT	0.82	CPT	0.79	CPT	0.82
IPT	0.68	CT	0.77	CT	0.79
PE	0.64	PE	0.63	PE	0.65
SIT	0.59	NET	0.49	PCT	0.40
NET	0.41	REL	0.38	NET	0.39
MBI	0.36	SIT	0.38	REL	0.39
PCT	0.34	PCT	0.36	MBI	0.38
REL	0.33	MBI	0.35	SIT	0.30
PsEd	0.09	PsEd	0.07	WL	0.09
WL	0.03	WL	0.07	PsEd	0.06

Note: EMDR = eye movement desensitization and reprocessing; CPT=cognitive processing therapy; CT= cognitive therapy; PE= prolonged exposure; NET= narrative exposure therapy; SIT= stress inoculation therapy; PsEd= psychoeducation; IPT= interpersonal therapy; MBI= mindfulness-based interventions; PCT= present – centered therapy; REL= relaxation; WL= waitlist

Appendix L2

Forest plot of all interventions compared to WL after the exclusion of outliers at the level of

pairwise meta-analyses.

Contrast to Waitlist	Comparison: other vs 'WL' (Random Effects Model)	SMD	95%-CI
CPT		-1.15	[-1.46; -0.84]
CT		-1.11	[-1.39; -0.83]
EMDR		-1.17	[-1.44; -0.90]
IPT		-0.97	[-1.62; -0.32]
MBI		-0.61	[-0.87; -0.35]
NET		-0.66	[-1.05; -0.28]
PCT		-0.58	[-0.89; -0.27]
PE		-0.91	[-1.11; -0.70]
PsEd		-0.14	[-0.57; 0.30]
REL		-0.57	[-0.93; -0.22]
SIT		-0.86	[-1.55; -0.17]
WL		0.00	
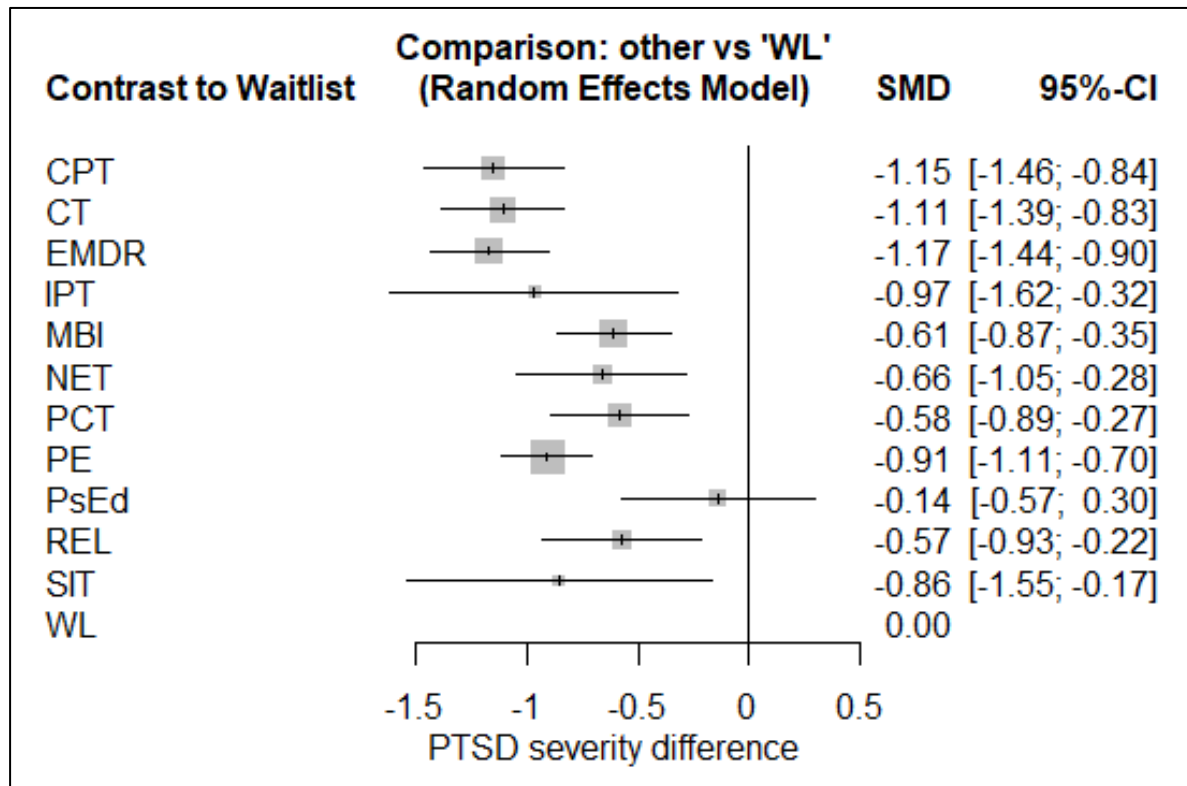
Appendix L2

Forest plot of all interventions compared to WL after the exclusion of outliers at network

level.

Contrast to Waitlist	Comparison: other vs 'WL' (Random Effects Model)	SMD	95%-CI
CPT		-1.17	[-1.51; -0.82]
CT		-1.13	[-1.44; -0.82]
EMDR		-1.36	[-1.66; -1.05]
IPT		-1.44	[-2.78; -0.10]
MBI		-0.58	[-0.87; -0.28]
NET		-0.81	[-1.21; -0.40]
PCT		-0.61	[-0.95; -0.26]
PE		-0.98	[-1.23; -0.73]
PsEd		0.04	[-0.55; 0.63]
REL		-0.62	[-1.05; -0.19]
SIT		-0.94	[-1.68; -0.20]
WL		0.00	

PTSD severity difference
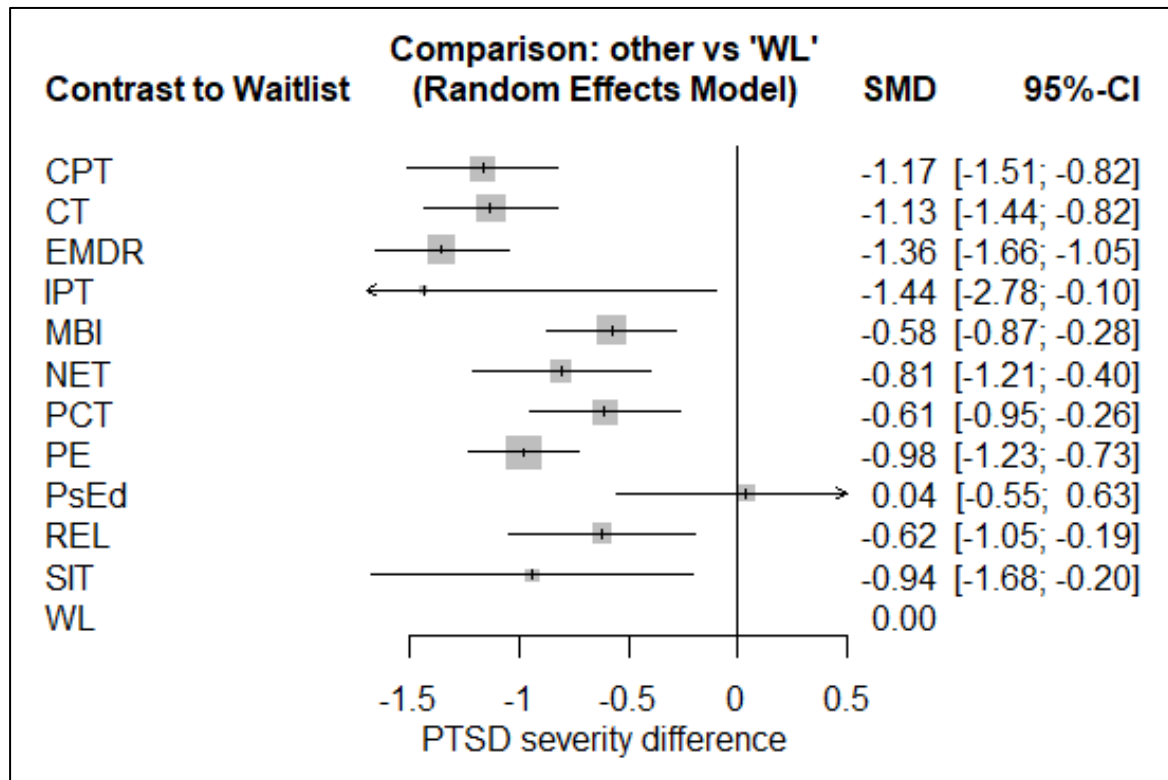
Appendix L4

Forest plot of all interventions compared to WL after the exclusion of all outliers

Contrast to Waitlist	Comparison: other vs 'WL' (Random Effects Model)	SMD	95%-CI
CPT		-1.16	[-1.47; -0.85]
CT		-1.12	[-1.40; -0.84]
EMDR		-1.25	[-1.54; -0.96]
IPT		-1.44	[-2.72; -0.16]
MBI		-0.57	[-0.84; -0.30]
NET		-0.63	[-1.02; -0.25]
PCT		-0.60	[-0.91; -0.28]
PE		-0.95	[-1.18; -0.73]
PsEd		0.10	[-0.44; 0.64]
REL		-0.57	[-0.97; -0.18]
SIT		-0.86	[-1.56; -0.17]
WL		0.00	

PTSD severity difference

Estimated effects of network meta-analysis after the exclusion of outliers at the level of pairwise meta-analyses

CPT											
-0.04 [-0.43; 0.35]	CT										
0.02 [-0.38; 0.43]	0.06 [-0.32; 0.44]	EMDR									
-0.18 [-0.90; 0.53]	-0.14 [-0.84; 0.55]	-0.20 [-0.89; 0.48]	IPT								
-0.54 [-0.94; -0.14]	-0.50 [-0.87; -0.13]	-0.56 [-0.92; -0.20]	-0.36 [-1.05; 0.33]	MBI							
-0.49 [-0.98; 0.00]	-0.45 [-0.92; 0.02]	-0.51 [-0.98; -0.05]	-0.31 [-1.06; 0.44]	0.05 [-0.40; 0.50]	NET						
-0.57 [-0.97; -0.17]	-0.53 [-0.93; -0.13]	-0.59 [-0.99; -0.19]	-0.39 [-1.09; 0.32]	-0.03 [-0.40; 0.34]	-0.08 [-0.57; 0.41]	PCT					
-0.24 [-0.59; 0.11]	-0.20 [-0.52; 0.12]	-0.26 [-0.57; 0.04]	-0.06 [-0.71; 0.59]	0.30 [-0.01; 0.60]	0.25 [-0.18; 0.67]	0.33 [0.02; 0.64]	PE				
-1.01 [-1.54; -0.48]	-0.97 [-1.49; -0.46]	-1.04 [-1.54; -0.53]	-0.83 [-1.60; -0.06]	-0.47 [-0.92; -0.03]	-0.52 [-1.06; 0.01]	-0.45 [-0.96; 0.07]	-0.77 [-1.23; -0.32]	PsEd			
-0.58 [-1.04; -0.11]	-0.54 [-0.96; -0.11]	-0.60 [-1.00; -0.20]	-0.40 [-1.06; 0.27]	-0.04 [-0.45; 0.37]	-0.09 [-0.61; 0.43]	-0.01 [-0.46; 0.44]	-0.34 [-0.71; 0.03]	0.44 [-0.11; 0.98]	REL		
-0.29 [-1.04; 0.46]	-0.25 [-0.99; 0.49]	-0.31 [-1.05; 0.42]	-0.11 [-1.05; 0.83]	0.25 [-0.48; 0.98]	0.20 [-0.52; 0.91]	0.28 [-0.47; 1.02]	-0.05 [-0.74; 0.64]	0.72 [-0.08; 1.52]	0.28 [-0.48; 1.05]	SIT	
-1.15 [-1.46; -0.84]	-1.11 [-1.39; -0.83]	-1.17 [-1.44; -0.90]	-0.97 [-1.62; -0.32]	-0.61 [-0.87; -0.35]	-0.66 [-1.05; -0.28]	-0.58 [-0.89; -0.27]	-0.91 [-1.11; -0.70]	-0.14 [-0.57; 0.30]	-0.57 [-0.93; -0.22]	-0.86 [-1.55; -0.17]	WL

Note: EMDR = eye movement desensitization and reprocessing; CPT=cognitive processing therapy; CT= cognitive therapy; PE= prolonged exposure; NET= narrative exposure therapy; SIT= stress inoculation therapy; PsEd= psychoeducation; IPT= interpersonal therapy; MBI= mindfulness-based interventions; PCT= present – centered therapy; REL= relaxation; WL= waitlist
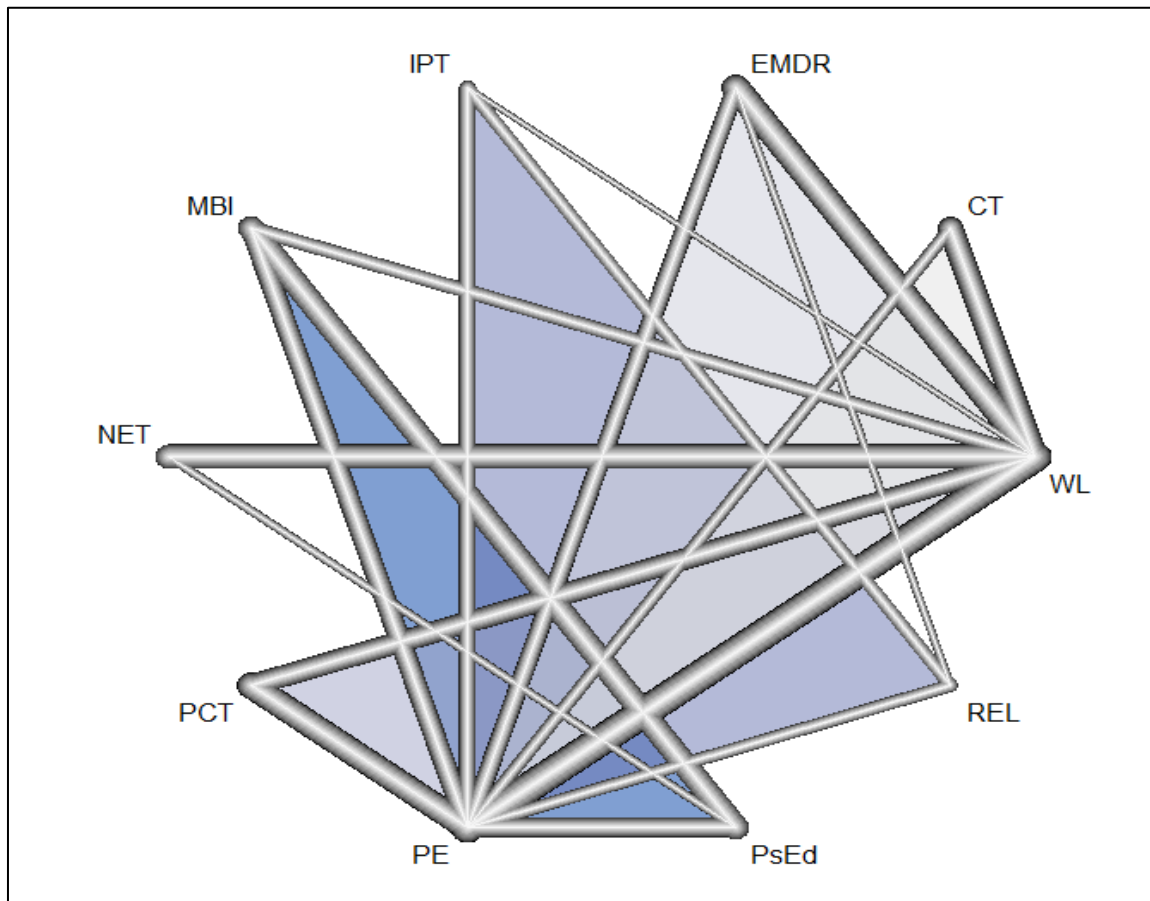
Appendix M2

Estimated effects of network meta-analysis after the exclusion of network-level outliers

CPT											
-0.04 [-0.46; 0.39]	CT										
0.19 [-0.27; 0.64]	0.23 [-0.20; 0.65]	EMDR									
0.27 [-1.11; 1.65]	0.31 [-1.06; 1.68]	0.08 [-1.29; 1.46]	IPT								
-0.59 [-1.04; -0.14]	-0.55 [-0.97; -0.14]	-0.78 [-1.20; -0.36]	-0.86 [-2.23; 0.51]	MBI							
-0.36 [-0.89; 0.17]	-0.32 [-0.83; 0.18]	-0.55 [-1.06; -0.04]	-0.63 [-2.03; 0.77]	0.23 [-0.26; 0.72]	NET						
-0.56 [-1.01; -0.11]	-0.52 [-0.97; -0.08]	-0.75 [-1.20; -0.30]	-0.83 [-2.21; 0.55]	0.03 [-0.39; 0.45]	-0.20 [-0.73; 0.33]	PCT					
-0.19 [-0.59; 0.21]	-0.15 [-0.52; 0.21]	-0.38 [-0.74; -0.01]	-0.46 [-1.82; 0.90]	0.40 [0.03; 0.77]	0.17 [-0.30; 0.64]	0.37 [0.02; 0.72]	PE				
-1.21 [-1.89; -0.53]	-1.17 [-1.84; -0.51]	-1.40 [-2.06; -0.74]	-1.48 [-2.94; -0.02]	-0.62 [-1.23; -0.01]	-0.85 [-1.50; -0.20]	-0.65 [-1.32; 0.03]	-1.02 [-1.66; -0.39]	PsEd			
-0.55 [-1.09; -0.01]	-0.51 [-1.00; -0.02]	-0.74 [-1.20; -0.27]	-0.82 [-2.22; 0.59]	0.04 [-0.44; 0.53]	-0.19 [-0.77; 0.40]	0.01 [-0.52; 0.54]	-0.36 [-0.82; 0.10]	0.66 [-0.06; 1.38]	REL		
-0.23 [-1.04; 0.58]	-0.19 [-0.98; 0.61]	-0.41 [-1.21; 0.38]	-0.50 [-2.03; 1.03]	0.36 [-0.43; 1.16]	0.13 [-0.63; 0.90]	0.33 [-0.47; 1.14]	-0.04 [-0.79; 0.71]	0.98 [0.06; 1.91]	0.32 [-0.53; 1.17]	SIT	
-1.17 [-1.51; -0.82]	-1.13 [-1.44; -0.82]	-1.36 [-1.66; -1.05]	-1.44 [-2.78; -0.10]	-0.58 [-0.87; -0.28]	-0.81 [-1.21; -0.40]	-0.61 [-0.95; -0.26]	-0.98 [-1.23; -0.73]	0.04 [-0.55; 0.63]	-0.62 [-1.05; -0.19]	-0.94 [-1.68; -0.20]	WL

Note: EMDR = eye movement desensitization and reprocessing; CPT=cognitive processing therapy; CT= cognitive therapy; PE= prolonged exposure; NET= narrative exposure therapy; SIT= stress inoculation therapy; PsEd= psychoeducation; IPT= interpersonal therapy; MBI= mindfulness-based interventions; PCT= present – centered therapy; REL= relaxation; WL= waitlist
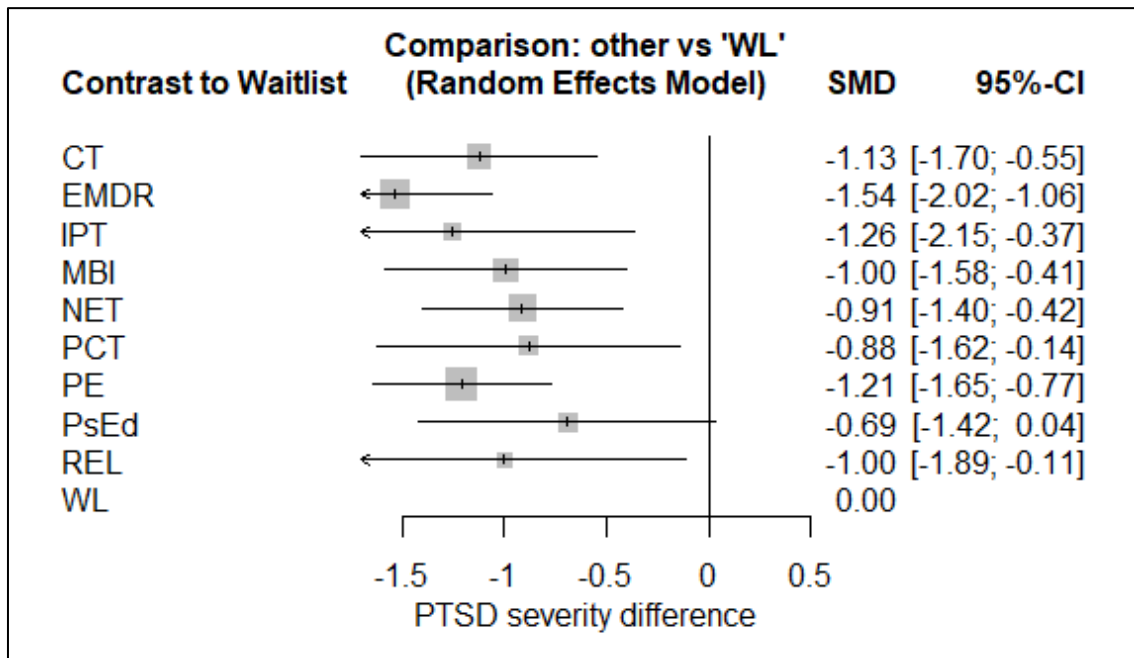
Appendix M3

Estimated effects of network meta-analysis after the exclusion of all outliers

CPT	CT	EMDR	IPT	MBI	NET	PCT	PE	PsEd	REL	SIT	WL
CPT											
-0.04 [-0.43; 0.35]	**CT**										
0.09 [-0.33; 0.51]	0.13 [-0.26; 0.53]	**EMDR**									
0.28 [-1.04; 1.60]	0.32 [-0.99; 1.63]	0.19 [-1.13; 1.50]	**IPT**								
-0.59 [-1.00; -0.18]	-0.55 [-0.94; -0.17]	-0.69 [-1.07; -0.30]	-0.87 [-2.18; 0.44]	**MBI**							
-0.53 [-1.02; -0.03]	-0.49 [-0.96; -0.01]	-0.62 [-1.10; -0.14]	-0.81 [-2.15; 0.53]	0.07 [-0.40; 0.53]	**NET**						
-0.56 [-0.97; -0.16]	-0.52 [-0.93; -0.12]	-0.65 [-1.07; -0.24]	-0.84 [-2.16; 0.48]	0.03 [-0.35; 0.41]	-0.04 [-0.53; 0.46]	**PCT**					
-0.20 [-0.56; 0.16]	-0.16 [-0.50; 0.17]	-0.30 [-0.64; 0.05]	-0.49 [-1.79; 0.82]	0.39 [0.05; 0.73]	0.32 [-0.12; 0.76]	0.36 [0.04; 0.67]	**PE**				
-1.26 [-1.88; -0.63]	-1.22 [-1.83; -0.61]	-1.35 [-1.97; -0.74]	-1.54 [-2.93; -0.15]	-0.67 [-1.23; -0.11]	-0.73 [-1.33; -0.13]	-0.70 [-1.31; -0.08]	-1.05 [-1.64; -0.47]	**PsEd**			
-0.58 [-1.08; -0.09]	-0.54 [-1.00; -0.09]	-0.68 [-1.11; -0.24]	-0.86 [-2.21; 0.48]	0.01 [-0.44; 0.45]	-0.06 [-0.61; 0.49]	-0.02 [-0.51; 0.46]	-0.38 [-0.80; 0.04]	0.68 [0.01; 1.34]	**REL**		
-0.29 [-1.05; 0.46]	-0.25 [-1.00; 0.49]	-0.39 [-1.13; 0.36]	-0.58 [-2.03; 0.88]	0.30 [-0.44; 1.04]	0.23 [-0.49; 0.95]	0.27 [-0.48; 1.01]	-0.09 [-0.79; 0.61]	0.97 [0.10; 1.83]	0.29 [-0.50; 1.08]	**SIT**	
-1.16 [-1.47; -0.85]	-1.12 [-1.40; -0.84]	-1.25 [-1.54; -0.96]	-1.44 [-2.72; -0.16]	-0.57 [-0.84; -0.30]	-0.63 [-1.02; -0.25]	-0.60 [-0.91; -0.28]	-0.95 [-1.18; -0.73]	0.10 [-0.44; 0.64]	-0.57 [-0.97; -0.18]	-0.86 [-1.56; -0.17]	**WL**

Note: EMDR = eye movement desensitization and reprocessing; CPT=cognitive processing therapy; CT= cognitive therapy; PE= prolonged exposure; NET= narrative exposure therapy; SIT= stress inoculation therapy; PsEd= psychoeducation; IPT= interpersonal therapy; MBI= mindfulness-based interventions; PCT= present – centered therapy; REL= relaxation; WL= waitlist

Appendix N1

Network graph of the NMA of studies with either an unclear risk of bias or high

primary study quality

Appendix N2

Forest plot of the NMA including only studies with either an unclear risk of bias or high

primary study quality

Contrast to Waitlist	Comparison: other vs 'WL' (Random Effects Model)	SMD	95%-CI
CT		-1.13	[-1.70; -0.55]
EMDR		-1.54	[-2.02; -1.06]
IPT		-1.26	[-2.15; -0.37]
MBI		-1.00	[-1.58; -0.41]
NET		-0.91	[-1.40; -0.42]
PCT		-0.88	[-1.62; -0.14]
PE		-1.21	[-1.65; -0.77]
PsEd		-0.69	[-1.42; 0.04]
REL		-1.00	[-1.89; -0.11]
WL		0.00	

PTSD severity difference

Appendix N3

Heatmap plot displaying hotspots of heterogeneity and inconsistency in the NMA of studies

with an unclear RoB or a high PSQ
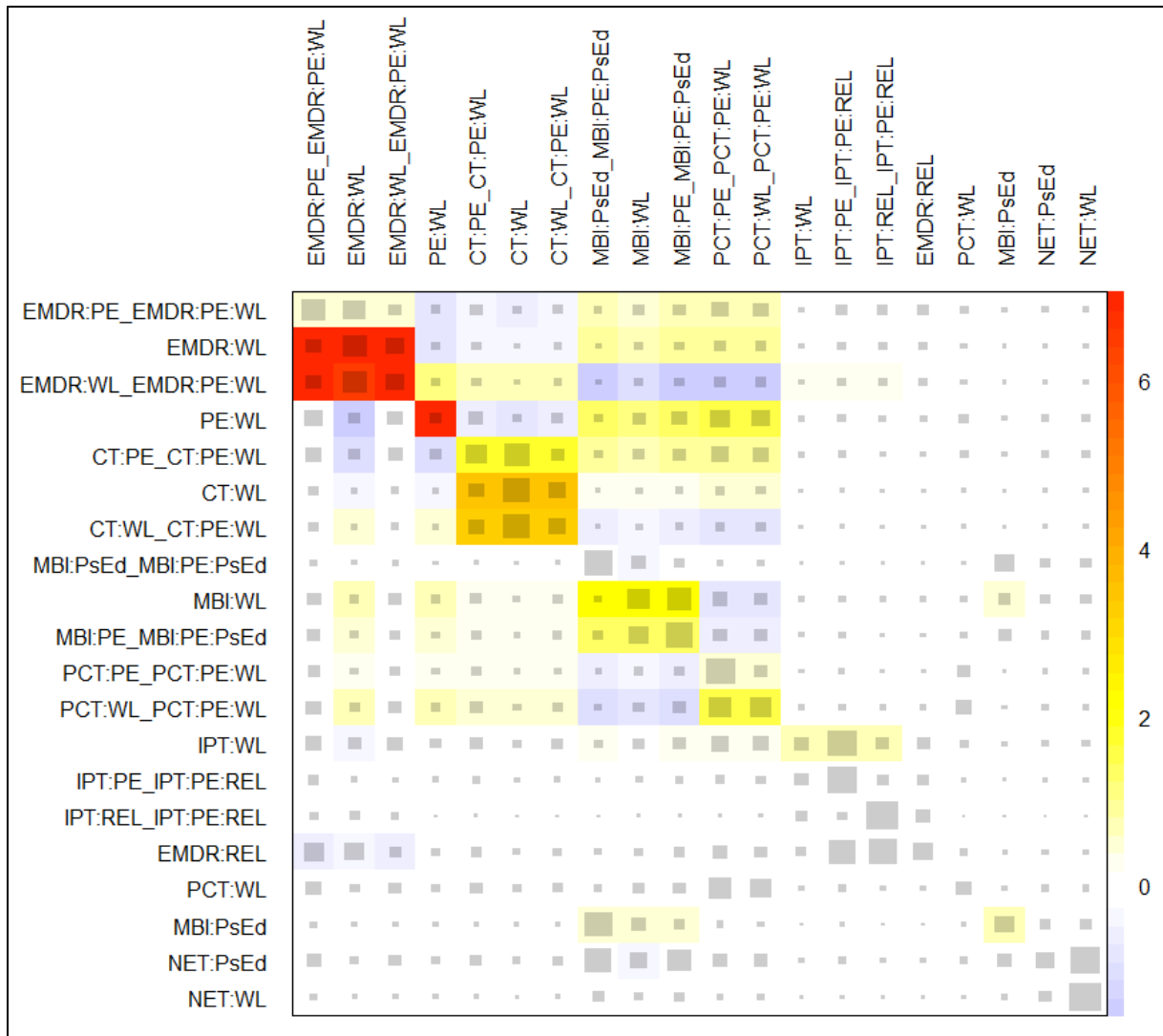
Appendix N4

NMA of studies with an unclear risk of bias/ high PSQ: P-score based treatment ranking

intervention	p-score
EMDR	0.90
IPT	0.68
PE	0.68
CT	0.60
MBI	0.50
REL	0.50
NET	0.43
PCT	0.42
PsEd	0.28
WL	0.01

Note: EMDR = eye movement desensitization and reprocessing; CT= cognitive therapy; PE= prolonged exposure; NET= narrative exposure therapy; PsEd= psychoeducation; IPT= interpersonal therapy; MBI= mindfulness-based interventions; PCT= present – centered therapy; REL= relaxation; WL= waitlist

Appendix O1

Summary of risk of bias evaluations of effect sizes based on intention-to-treat analysis

Unique ID	2	Study ID	20201	Assessor	R
Ref or Label	Acarturk 2016	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	EMDR	Comparator	WL	Source	Journal article(s) with results of the trial; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
Outcome	self-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "After including the participants, another researcher, not involved in the current study, used a computer generated random-number list for the allocation of participants to different treatment groups. Participants were randomly assigned on a 1:1 basis to the EMDR or wait-list group."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote. "The average age of the EMDR group was 33.32 years (range=18–59 years), and of the control group 34.04 years (range=17–64 years). The groups also did not differ on their scores on BDI-II, IES-R,HTQ and HSCL total and HSCL depression scales."
	Risk of bias judgement	**Some concerns**	computer generated random-number list , unclear allocation concealment
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	Y	Quote: "The participants and the therapists were necessarily aware of the allocated arm, but the outcome assessors were kept blind to the allocation."
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	

	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	N	comment: No such deviations reported.
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	PY	Quote: "All outcome analyses were conducted according to the intention-to-treat principle. We used linear mixed models to analyse changes over time [...]"
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no deviations reported; ITT analysis (linear mixed models)
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "A total of 98 participants were randomly assigned either to receive EMDR therapy (n=49) or to be wait-listed (n=49) as the control group. In all, 37 and 33 people remained in the respective EMDR and control groups for the post-test assessments. "
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	Quote: "Drop-outs did not differ significantly from the completers [...]"
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NI	Quote: "Reasons for drop-out in the EMDR group were refusal of entering into treatment (n=7) and moving out of the camp (n=5). reasons of drop-out in the control group were moving out of the camp (n=9) and refusal (n=7)."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	comment: proportions of dropout =24.48% in WL, = 32.65% in EMDR (no significance test of difference); The number and proportion of dropouts due to refusal were balaced across both groups (7 out of 49), suggesting that the propability of refusal was not related to either condition. However, the exact reasons for refusal are not clear from this information, so a dependency of missingness on its true value cannot be ruled out. Dropout due to

			removal from the refugee camp, however, is likely to be unrelated to the treatment. Additionally, Participants with missing data at t2 did not differ significantly from completers.
	Risk of bias judgement	**Some concerns**	It is not likely that missingness in the outcome depends on its true value (see description). However, there are concerns.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The IES-R is a 22-item self-report instrument which rates the severity of PTSD symptoms (Weiss & Marmar, 1997)." "The test–retest reliability calculated by administering the scale to the same sample on two occasions, 2 weeks apart, yielded a Pearson correlation coefficient of r=0.88 (M Zaghrout, unpublished observations). In the present study, the baseline administration of the scale yielded a Cronbach's α value of 0.87, indicating a good internal consistency of all items."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	comment: the 'assessors' were -in this case- the participants themselves (who were aware of their intervention).
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: assessment could have been influenced because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NI	comment: the risk of bias due to knowledge of the intervention is high, as the comparator is a passive control condition (here: WL). It would be lower if the comparator was another active intervention.
	Risk of bias judgement	**High**	comment: the risk of bias due to knowledge of the intervention is rated high because the comparator is a no-treatment condition (here: WL)

Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: "The study was registered to Clinical Trials (ClinicalTrials.gov identifier: NCT01847742). [...] The Consolidated Standards of Reporting Trials (CONSORT)checklist is available as supporting information." comment: Registration and history of changes examined and compared to the report; information is consistent with reported results.
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest (IES, HTQ)
	5.3 ... multiple eligible analyses of the data?	PN	comment: mITT data based on LMMs is available;
	Risk of bias judgement	**Low**	Trial Registration/ history of changes examined and compared to the report; information is consistent with reported results.
Overall bias	**Risk of bias judgement**	**High**	unclear allocation concealment. ITT analysis (linear mixed models). high risk of bias due to participants' knowledge of the intervention. Trial Registration/ history of changes examined and compared to the report; information is consistent with reported results.

Unique ID	6	**Study ID**	60101	**Assessor**	R
Ref or Label	Bolton 2014	**Aim**	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	CPT	**Comparator**	WL	**Source**	Journal article(s) with results of the trial

Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "Randomization of CMHWs and participant IDs was done using JB Stata's randomization function." "The CMHWs received twenty participant IDs randomly assigned to intervention or control [...]" "if a person consented, the CMHW opened a sealed envelope attachedto the consent form containing theparticipant's assignment"
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	comment: see table 3
	Risk of bias judgement	**Low**	allocation concealment using sealed envelopes; no substantial baseline differences but clinical scores slightly higher in CPT group (no sign. test).
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	Y	comment: Participants must necessarily have been aware of their assigned condition because WL participants were informed that they would receive treatment later
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "All analyses were conducted on the full intent to treat sample" "Multiple imputation by chained equations accounted for missing scale items and follow up scores among those lost to follow up"

	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding, no deviations, ITT analysis using MI.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	comment: considerable dropout rate: see Figure 1 flow-chart
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NI	Quote: "Participants who dropped out of the trial after having started the treatment rarely gave reasons beyond not wanting to continue. One CPT participant moved away, one was referred for psychosis, and one left after being verbally abused by her husband for
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	getting treatment. This was the only significant harm or unintended effect reported in the study. One control was referred to a psychiatrist for worsening symptoms."
			"Those who did not begin or dropped out of CPT were more likely to be male and married compared with those who completed treatment."
			comment: no difference between groups in the proportions of non-starters/ dropouts, reported reasons for dropouts did not differ substantially, however, most gave no reason beyond not wanting to continue so it is possible that reasons were related to the treatment condition
	Risk of bias judgement	**High**	considerable dropout rate; no difference between groups with respect to the proportions of dropouts; most gave no reason beyond not wanting to continue; participants with/without posttest data differed on two variables; all in all, risk of bias is high.

Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	PN	Quote: "Qualitative study data were used to adapt the Hopkins Symptom Checklist for Depression and Anxiety (HSCL-25) [18,19], the Harvard Trauma Questionnaire (HTQ) [20], and the Inventory of Traumatic Grief [21,22] to measure symptoms of depression, anxiety, posttraumatic stress and traumatic grief. Adaptation included adding 13 locally relevant symptoms." comment: the HTQ is a validated PTSD scale and likely to be sensitive to treatment effects. Adaptations are reported in detail, as are justifications and the rationale behind
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "CMHWs or supervisors blind to participants' treatment status did 197 (85%) of the interviews; 35 (15%) were implemented by un-blinded CMHWs. The latter group included participants who terminated treatment and refused further contact. Rather than forgo assessment, the treating CMHW did the interview." comment: beyond the limited blindness of interviewers, there was no blind assessment: Accoring to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: the assessors were no independent researchers but involved in the study, which would otherwise lower the risk of bias.

	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NI	Also, the risk of bias due to knowledge of the intervention is rated higher if the comparator is no treatment (here: WL) than when the comparator is another active intervention. In addition, for subjective outcomes such as 'clinical impression of improvement', knowledge of the intervention received could be highly influential.
	Risk of bias judgement	**High**	no blindness of assessors, passive control condition and subjective outcome which is especially sensitive to bias due to knowledge of the intervention
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: "Trial Registration: ClinicalTrials.gov NCT00925262" comment: History of changes was examined. information consistent with information in the report
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: ITT data is available for all participants; generally the risk of bias due to mutiple eligible analyses of the data is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Low**	Trial was registered. History of changes was examined. information consistent with information in the report
Overall bias	**Risk of bias judgement**	**High**	no blindness of assessors, passive control condition and subjective outcome which is especially sensitive to bias due to knowledge of the intervention. considerable dropout rate; no difference between groups with respect to the proportions of dropouts; most gave no reason beyond not wanting to continue; participants with/without posttest data differed on two variables; all in all, risk of bias is high. Trial was registered. History of changes was examined. information consistent with information in the report

Unique ID	10	Study ID	80102	Assessor	R	
Ref or Label	Bormann 2018	**Aim**	assignment to intervention (the 'intention-to-treat' effect)			
Experimental	MBI	**Comparator**	PCT	**Source**	Journal article(s) with results of the trial	
Outcome	clinician-rated PTSD	**Results**	SMD(between)	**Weight**	1	

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "All participants provided written informed consent and were randomly assigned to treatment arms by study coordinators using sealed lists of computer-generated random numbers from the study statistician."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "We detected no significant differences between the mantram and present-centered therapy groups at baseline on any demographic characteristics, medication use, or clinical measures (Table 1) or regarding the credibility or expectationsconcerningthetreatment(seeAppendix2intheonline supplement)."
	Risk of bias judgement	**Low**	sealed lists of computer-generated random numbers; no significant baseline differences.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were necessarily aware of the assigned condition
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	

	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	comment: ITT analysis with linear mixed models
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding of participants/therapists; no deviations; ITT analysis.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	comment: 20 missings in MBI group and 12 in PCT group at post-assessment
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	comment: no analysis correcting for bias or sensitivity analysis
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	comment: proportions of dropouts between groups =22.47% in MBI and =14% in PCT; statistical significance not reported. Documented reasons of dropout suggest that missingness in the outcome might depend on its true value in some cases (MBI n=7 dropped out for reasons that might be related to treatment; PCT n= 7 such cases). For the MBI group these cases represent 7.8% of participants, for the PCT group 8.3% (again, equal proportions).
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PY	
	Risk of bias judgement	**High**	proportions of dropouts between groups =22.4% in MBI and =14% in PCT; statistical significance not reported. Documented reasons of dropout suggest that missingness in the outcome might depend on its true value in some cases.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "CAPS scores can range from zero to 136, with higher scores indicating greater severity (19). A reduction of $10 points is considered a clinically meaningful improvement (17). Cronbach's alpha was 0.91." comment: The administered scale (CAPS) is a validated, gold-standard PTSD measure and likely to be sensitive to intervention effects

	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "blinded assessment before treatment, after treatment" comment: although the assessors were blind to the participants' condition, there was no blind assessment: According to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention. Thus, the risk that participants have answered according to their beliefs/expectations about the intervention effect is not as high (as opposed to a passive control condition)
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN	
	Risk of bias judgement	**Some concerns**	unblinded assessment; the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	Quote: "For all outcomes, we computed means and confidence intervals." comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: both unadjusted and adjusted outcome results are reported
	Risk of bias judgement	**Some concerns**	

Overall bias	Risk of bias judgement			High	unblinded assessment; the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention. proportions of dropouts between groups =22.4% in MBI and =14% in PCT; statistical significance not reported. Documented reasons of dropout suggest that missingness in the outcome might depend on its true value in some cases

Unique ID	11	Study ID	110101	Assessor	R
Ref or Label	Bränström 2010	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	MBI	Comparator	WL	Source	Journal article(s) with results of the trial
Outcome	self-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "Random selection of participants to either the intervention or control group was done consecutively using a random sequence of numbers indicating group assignment. Once a participant was recruited to the study, he/she was assigned a study number and was assigned to the intervention or control group according to the sequence of numbers. The sequence was produced through the SPSS software's random selection procedure."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	"There were no significant differences between the intervention and control group concerning age [..], education [...], work status [...], household income [...], or use of antidepressants [...]. In addition, no differences were found in any of the psychological outcome variables—perceived stress, depression, anxiety,

			positive states of mind, posttraumatic stress symptoms, or mindfulness - indicating that the randomization was successful."
	Risk of bias judgement	**Some concerns**	sequence was produced through the SPSS software's random selection procedure; unclear allocation concealment; no baseline differences.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	Y	Quote: "No blinding of group assignment was done."
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	comment: therapists and participants were necessarily aware of the assigned condition
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	PY	Quote: " An initial intention-to-treat analysis was conducted with missing data at follow-up imputed according to last-observation-carriedforward strategy."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding of participants/therapists; no deviations; ITT analysis.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	comment: 28 participants dropped out
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NI	

	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	comment: 14 participants dropped out before baseline assessment. there were equal proportions of dropouts (7 in MBI, 7 in WL). proportions are also equal regarding dropouts after baseline assessment (n=7 in the MBI group, and n=7 in WL). However, reasons for dropout are not reported for all cases and not in detail, thus raising some concerns. However, considering equal proportions and the explicit report that no adverse effects or side effects occured it is not very likely that missingness in the outcome depended on its true value.
	Risk of bias judgement	Some concerns	equal proportions of dropouts; reasons for dropout are not reported for all cases; there are concerns.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	PN	comment: The administered scale (IES) is a validated PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "Questionnaires were sent to the patients by mail directly after randomization and at 3 and 6 months after randomization, together with a prepaid return envelope." comment: comment: the 'assessors' were -in this case- the participants themselves (who were aware of their intervention). Accoring to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is rated higher if the comparator is no treatment (here: WL) than when the comparator is another active intervention.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	

	Risk of bias judgement	High	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	
	5.3 ... multiple eligible analyses of the data?	PN	
	Risk of bias judgement	Some concerns	
Overall bias	Risk of bias judgement	High	non-adherence (not adressed here) in the intervention group (participants attended a different number of sessions) leads to a risk of bias towards the null, in other words, to a risk of underestimating the PP-effect of the MBI intervention in reducing PTSD symptoms. The lack of assessment of treatment fidelity constitutes another risk of bias. There is an additional risk of bias (favoring the experimental) due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect (risk is especially relevant because comparator is a passive control condition).

Unique ID	14	**Study ID**	150101	**Assessor**	R
Ref or Label	Bryant 2008	**Aim**	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	PE	**Comparator**	CT	**Source**	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "Participants were informed that they would be randomly allocated to 1 of 3 treatment conditions. Randomization was conducted by a process of minimization stratified by sex, trauma type, and Acute Stress Disorder Interview score. Participants were assigned to groups using a random numbers system administered by an individual who worked at a site that was distant from the treatment center and was not otherwise involved with the study."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	NI	comment: no significance test reported, but descriptively PTSD severity at baseline differs between groups [PE =70.6 (17.7), CT= 66.8 (19.0), WL= 63.6 (18.3)]
	Risk of bias judgement	**Some concerns**	randomization by minimization; no information on allocation concealment and insufficient information regarding baseline characteristics.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	NI	comment: each therapist must have necessarily been aware of the participant's treatment condition during treatment; therapists were not aware of other participants' assigned condition if they were treated by other therapists in the study
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	Quote: "Individual therapy was conducted by 1 of 6 experienced master's degree–level clinical psychologists (J.M.,K.L.F.,S.H.,L.K.,E.K., or C.C.)who were trained to use treatment manuals and who received weekly supervision (R.A.B.). All therapists provided each type of treatment." ".Posttreatment and 6-month follow up assessments were conducted by independent clinical psychologists (J.M.,K.L.F.,S.H.,L.K.,E.K.,and C.C.)who were unaware of the participants' treatment groups."
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	N	comment: see Figure 1; no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol that might lead to bias

	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "We report completer analyses and intent-to-treat analyses, in which we used the last observation carried forward procedure."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding of participants/therapists; no deviations; ITT analysis.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	see Figure 1 flow-chart comment: PE post-assessment of 25 out of 30 randomized participants; CT post-assessment of 23 out of 30 randomized participants
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	Quote: "Planned comparisons of treatment completers and treatment dropouts indicated no differences between those who did and did not drop out of treatment on any pretreatment psychopathological or demographic variables." "9 there was no difference in drop-out rates for the PE and CR groups (17% vs 23%)."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	comment: documented and reported reasons for dropout are likely to be unrelated to the treatment condition (see Figure 1)

	Risk of bias judgement	Some concerns	documented and reported reasons for dropout are likely to be unrelated to the treatment condition; no differences between those who did and did not drop out of treatment; no difference in drop-out rates.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "Clinical psychologists diagnosed ASD using the Acute Stress Disorder Interview." "Initial assessments were conducted at the pretreatment session,before randomization. Posttreatment and 6-month followup assessments were conducted by independent clinical psychologists (J.M.,K.L.F.,S.H.,L.K.,E.K.,and C.C.) who were unaware of the participants' treatment groups."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "[...] assessments were conducted by independent clinical psychologists (J.M.,K.L.F.,S.H.,L.K.,E.K.,and C.C.) who were unaware of the participants' treatment groups."
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "Blindness was maintained by ensuring that clinicians who conducted assessments did not have access to participants' medical records or group allocation." comment: although the assessors were blind to the participants' condition, there was no blind assessment: Accoring to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention. Thus, the risk that participants have answered according to their beliefs/expectations about the intervention effect is not as high (as opposed to a passive control condition).
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN	

	Risk of bias judgement	Some concerns	no guarantee for complete blindness in view of the fact that participants might have been aware of their treatment condition. However, the risk of bias is lowered by the fact that the comparator was also an active treatment condition (plus same number of sessions, same session duration and both cognitive-behavioral interventions).
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	comment: results were reported for all PTSD measures assessed and for both time points (Table 2)
	5.3 ... multiple eligible analyses of the data?	PN	comment: both ITT and completers data is available; generally the risk of bias due to mutiple eligible analyses of the data is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	Some concerns	
Overall bias	Risk of bias judgement	Some concerns	randomization by minimization; no information on allocation concealment and baseline characteristics. there was no guarantee for complete blindness given the fact that participants might have been aware of their treatment condition. However, the risk of bias is lowered by the fact that the comparator was also an active treatment condition (plus same number of sessions, same session duration and both cognitive-behavioral interventions).

Unique ID	16	Study ID	150102	Assessor	R
Ref or Label	Bryant 2008	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	PE	Comparator	WL	Source	Journal article(s) with results of the trial

Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "Participants were informed that they would be randomly allocated to 1 of 3 treatment conditions. Randomization was conducted by a process of minimization stratified by sex, trauma type, and Acute Stress Disorder Interview score. Participants were assigned to groups using a random numbers system administered by an individual who worked at a site that was distant from the treatment center and was not otherwise involved with the study."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	NI	comment: no significance test reported, but descriptively PTSD severity at baseline differs between groups [PE =70.6 (17.7), CT= 66.8 (19.0), WL= 63.6 (18.3)]
	Risk of bias judgement	**Some concerns**	randomization by minimization; no information on allocation concealment and insufficient information regarding baseline characteristics.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	NI	comment: each therapist must have necessarily been aware of the participant's treatment condition during treatment; therapists were not aware of other participants' assigned condition if they were treated by other therapists in the study
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	Quote: "Individual therapy was conducted by 1 of 6 experienced master's degree–level clinical psychologists (J.M.,K.L.F.,S.H.,L.K.,E.K., or C.C.)who were trained to use treatment manuals and who received weekly supervision (R.A.B.). All therapists provided each type of treatment." ".Posttreatment and 6-month follow up assessments were conducted by independent clinical psychologists (J.M.,K.L.F.,S.H.,L.K.,E.K.,and C.C.)who were unaware of the participants' treatment groups."

	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	N	comment: see Figure 1; no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "We report completer analyses and intent-to-treat analyses, in which we used the last observation carried forward procedure."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding of participants/therapists; no deviations; ITT analysis.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	see Figure 1 flow-chart comment: PE post-assessment of 25 out of 30 randomized participants; WL post-assessment of 21 out of 30 randomized participants
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	Quote: "Planned comparisons of treatment completers and treatment dropouts indicated no differences between those who did and did not drop out of treatment on any pretreatment psychopathological or demographic variables."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	"[PE:] 25 Completed treatment, 1 Declined to participate, 1 Commenced medication, 2 Had an adverse reaction, 1 Moved" "[WL:] 21 Completed the study, 6 Declined to participate, 3 Moved"

			comment: reasons for dropout might be related to the assigned treatment condition in a few cases; equal proportions of dropouts; participants with and without posttest data did not differ on any baseline or demographic variables
	Risk of bias judgement	**Some concerns**	reasons for dropout might be related to the assigned treatment condition in a few cases which causes some concerns; but no differences between those who did and did not drop out of treatment; no difference in drop-out rates between groups.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "Clinical psychologists diagnosed ASD using the Acute Stress Disorder Interview." "Initial assessments were conducted at the pretreatment session,before randomization. Posttreatment and 6-month followup assessments were conducted by independent clinical psychologists (J.M.,K.L.F.,S.H.,L.K.,E.K.,and C.C.) who were unaware of the participants' treatment groups."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "[...] assessments were conducted by independent clinical psychologists (J.M.,K.L.F.,S.H.,L.K.,E.K.,and C.C.) who were unaware of the participants' treatment groups."
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "Blindness was maintained by ensuring that clinicians who conducted assessments did not have access to participants' medical records or group allocation." comment: although the assessors were blind to the participants' condition, there was no blind assessment: Accoring to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: assessment could have been influenced because participants were not blind to their condition and might have

	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	answered according to their beliefs/expectations about the intervention effect. the risk of bias due to knowledge of the intervention is high, as the comparator is a passive control condition
	Risk of bias judgement	**High**	the risk of bias due to knowledge of the intervention is high as participants were not blinded and the comparator was a no-treatment control condition
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	comment: results were reported for all PTSD measures assessed and for both time points (Table 2)
	5.3 ... multiple eligible analyses of the data?	PN	comment: both ITT and completers data is available; generally the risk of bias due to mutiple eligible analyses of the data is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Some concerns**	
Overall bias	**Risk of bias judgement**	**High**	randomization by minimization; no information on allocation concealment and baseline characteristics. some concerns regarding missing outcome data as reasons for dropout might have been related to the assigned treatment in a few cases. the risk of bias due to knowledge of the intervention is high as participants were not blinded and the comparator was a no-treatment control condition

Unique ID	17	**Study ID**	150103	**Assessor**	R
Ref or Label	Bryant 2008	**Aim**	assignment to intervention (the 'intention-to-treat' effect)		

Experimental	CT	Comparator	WL	Source	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "Participants were informed that they would be randomly allocated to 1 of 3 treatment conditions. Randomization was conducted by a process of minimization stratified by sex, trauma type, and Acute Stress Disorder Interview score. Participants were assigned to groups using a random numbers system administered by an individual who worked at a site that was distant from the treatment center and was not otherwise involved with the study."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	NI	comment: no significance test reported, but descriptively PTSD severity at baseline differs between groups [PE =70.6 (17.7), CT= 66.8 (19.0), WL= 63.6 (18.3)]
	Risk of bias judgement	**Some concerns**	randomization by minimization; no information on allocation concealment and insufficient information regarding baseline characteristics.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	NI	comment: each therapist must have necessarily been aware of the participant's treatment condition during treatment; therapists were not aware of other participants' assigned condition if they were treated by other therapists in the study
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	Quote: "Individual therapy was conducted by 1 of 6 experienced master's degree–level clinical psychologists (J.M.,K.L.F.,S.H.,L.K.,E.K., or C.C.)who were trained to use treatment manuals and who received weekly supervision (R.A.B.). All therapists provided each type of treatment." ".Posttreatment and 6-month follow up assessments were conducted by independent clinical psychologists

			(J.M.,K.L.F.,S.H.,L.K.,E.K.,and C.C.)who were unaware of the participants' treatment groups."
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	N	comment: see Figure 1; no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "We report completer analyses and intent-to-treat analyses, in which we used the last observation carried forward procedure."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding of participants/therapists; no deviations; ITT analysis.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	see Figure 1 flow-chart comment: PE post-assessment of 25 out of 30 randomized participants; CT post-assessment of 23 out of 30 randomized participants
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	Quote: "Planned comparisons of treatment completers and treatment dropouts indicated no differences between those who did and did not drop out of treatment on any pretreatment psychopathological or demographic variables." "[CT:] 23 Completed treatment, 2 Declined to participate, 1 Commenced medication, 2 Had an adverse reaction, 2 Moved"
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	

			"[WL:] 21 Completed the study, 6 Declined to participate, 3 Moved"
			comment: reasons for dropout might be related to the assigned treatment condition in a few cases; equal proportions of dropouts; participants with and without posttest data did not differ on any baseline or demographic variables
	Risk of bias judgement	**Some concerns**	reasons for dropout might be related to the assigned treatment condition in a few cases which causes some concerns; but no differences between those who did and did not drop out of treatment; no difference in drop-out rates between groups.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "Clinical psychologists diagnosed ASD using the Acute Stress Disorder Interview." "Initial assessments were conducted at the pretreatment session,before randomization. Posttreatment and 6-month followup assessments were conducted by independent clinical psychologists (J.M.,K.L.F.,S.H.,L.K.,E.K.,and C.C.) who were unaware of the participants' treatment groups."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "[...] assessments were conducted by independent clinical psychologists (J.M.,K.L.F.,S.H.,L.K.,E.K.,and C.C.) who were unaware of the participants' treatment groups."
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "Blindness was maintained by ensuring that clinicians who conducted assessments did not have access to participants' medical records or group allocation." comment: although the assessors were blind to the participants' condition, there was no blind assessment: Accoring to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be

			considered to be aware of intervention received unless convincing evidence is available to the contrary.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. the risk of bias due to knowledge of the intervention is high, as the comparator is a passive control condition
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	Y	
	Risk of bias judgement	**High**	the risk of bias due to knowledge of the intervention is high as participants were not blinded and the comparator was a no-treatment control condition
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	comment: results were reported for all PTSD measures assessed and for both time points (Table 2)
	5.3 ... multiple eligible analyses of the data?	PN	comment: both ITT and completers data is available; generally the risk of bias due to mutiple eligible analyses of the data is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Some concerns**	
Overall bias	**Risk of bias judgement**	**High**	randomization by minimization; no information on allocation concealment and baseline characteristics. the risk of bias due to knowledge of the intervention is high as participants were not blinded and the comparator was a no-treatment control condition

Unique ID	19	Study ID	200102	Assessor	R	
Ref or Label	Cigrang 2017	Aim	assignment to intervention (the 'intention-to-treat' effect)			
Experimental	PE	Comparator	WL	Source	Journal article(s) with results of the trial; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)	
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1	

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: " Those who met eligibility criteria were block randomized using a customized web based application [...]"
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "The baseline demographic, symptom, and military-specific characteristics of the two groups were similar and no statistical differences were found (p > .05, see Table 1). In addition, the two groups were similar at time of study enrollment on percentage who reported taking a PTSD medication (19% vs. 18%) and the mean total number of medications (3.6 vs. 3.7)."
	Risk of bias judgement	Some concerns	block randomization; no information on allocation concealment; no substantial baseline differences.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists (and participants) were necessarily aware of the assigned condition
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias

	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "Analyses were conducted by the third author and used all data from all 67 participants randomized initially to PE-PC (N = 34) or MCC (N = 33) using an intent-to-treat model."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding of participants/therapists; no deviations; ITT analysis.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	see Figure 1 CONSORT comment: for WL group data for almost all participants available (31/33); for PE group more missing outcome data (available for n=26/34)
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	comment: no analysis correcting for bias or sensitivity analysis or the like reported
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: Flow-chart: "34 Allocated to PE-PC, 26 Posttreatment Analyzed: 2 Lost to Posttreatment: 2 Scheduling conflicts; 6 Did not receive full intervention: 4 requested to drop from treatment, 1 withdrawn by PI discretion, 1 lost to contact during treatment"
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	"33 Allocated to MCC, 31 Post-MCC Analyzed: 2 Lost to Post-MCC, 1 hospitalized, 1 requested to drop from study" comment: unequal proportions of dropouts between groups (WL 6%, PE 23.5%); documented reasons do not eliminate the possibility of a dependency of missingness in the outcome on its true value

		Risk of bias judgement	High	unequal proportions of dropouts between groups (WL 6%, PE 23.5%); documented reasons might be treatment-related.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?		PN	comment: The administered scale (PCL) is a validated PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?		PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?		PY	Quote: "Baseline, posttreatment/postminimal contact control, and 8-week and 6-month follow-up assessments using all measures were conducted by an independent evaluator who was trained to administer the study measures and was a member of the research team." comment: no explicit information on blindness of those assessors provided. Even if the assessors were blind to the participants' condition, there was no blind assessment: According to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?		PY	comment: The quote indicates that the assessor was no independent researcher not involved in the study, which would otherwise lower the risk of bias, but a member of the research team.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?		NI	
	Risk of bias judgement		**High**	comment: there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect (risk is especially relevant because comparator is a passive control condition). As the assessor was no independent researcher there is a risk of bias.

Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: "Trial Registration: Clinicaltrials.gov identifier: NCT02290639"
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	
	5.3 ... multiple eligible analyses of the data?	PN	comment: ITT data for all participants randomized are available; generally the risk of bias due to mutiple eligible analyses of the data is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**High**	

Unique ID	20	**Study ID**	220101	**Assessor**	R
Ref or Label	van den Berg 2015	**Aim**	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	PE	**Comparator**	EMDR	**Source**	Journal article(s) with results of the trial; Trial protocol; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
Outcome	clinician-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "An independent randomization bureau randomized the treatment condition using stratified randomization blocks per therapist with equal strata sizes.Therapists confirmed the treatment assignment in writing. Data were stored at the study coordination center."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	

	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: " At baseline, there were no significant differences between the groups in any of the demographic or clinical characteristics."
	Risk of bias judgement	**Low**	independent randomization bureau randomized the treatment condition using stratified randomization blocks, no significant differences between the groups in any of the demographic or clinical characteristics.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists (and participants) may necessarily have been aware of the assigned condition
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	N	see Figure 1 flow chartcomment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol that might lead to biasQuote: "After treatment and at the 6-month follow-up, patient files were reviewed to check whether trauma-focused treatments had taken place and if there had been any changes in the prescribed medications, as well as for any deviations from standard care.""Adherence to protocols was rated as good or excellent in 91.2%of PE sessions and 97.1% of EMDR sessions.""There were no differences between groups in additional support provided by caregivers. Groups did not differ in the percentage of participants receiving additional non–trauma focused psychotherapy during treatment (17.0% in PE, 20.8% in EMDR, and 21.3% in WL)"
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "Continuous variables were analyzed on an intent-to-treat basis with linear mixed models (LMMs). [...] Analyses of completers and intent-to-treat analyses with last observation carried forward (with missing data on loss of diagnosis conservatively replaced with a negative value [ie, no loss of diagnosis]) were performed to test the robustness of the findings."

	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding of participants/therapists; no deviations; ITT analysis.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	see Figure 1 flow-chart comment: post-assessment of 47 out of 53 participants in PE group, and of 44 out of 55 participants in EMDR group
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NI	Quote: "There was no difference in dropout between the PE (13 participants [24.5%]) and EMDR (11 participants [20.0%]) (P=.57)."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	Quote: "Sensitivity Analyses: Completer analyses were performed (n=113), among which no baseline differences were observed between groups in any of the demographic or clinical characteristics. In addition, intent-to-treat analyses with last observation carried forward were performed (n = 155). All results for the CAPS, PSS-SR, and PTCI were similar to the results from the intent-to-treat analyses, thereby underlining the robustness of the findings." "There was no difference in dropout between the PE (13 participants [24.5%])and EMDR (11 participants [20.0%]) (P=.57)." comment: the information in the quote on "sensitivity analyses"

			does not seem sufficient for the judgement that there is evidence that the result was not biased by missing outcome data (3.2). no information is provided on reasons for dropout. Missings might not be missing at random, calling for more complex methods for sensitivity analyses (e.g. selection models or pattern mixed models). Proportions of dropouts did not differ between the groups. However, given the fact that reasons for dropout are not reported, there are some concerns
	Risk of bias judgement	**Some concerns**	no difference in dropout; no information on reasons for dropout.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote from the published study protocol: "The CAPS is considered the gold standard to diagnose posttraumatic stress disorder as defined in the DSM-IV-TR and to establish its severity. A review of the empirical literature on psychometric properties of the CAPS [76] indicates that the CAPS has excellent reliability (>0.90), yielding consistent scores across items, raters and testing occasions. There is also strong evidence of validity: the CAPS has excellent (>0.90) convergent and discriminant validity, diagnostic utility, and sensitivity to clinical change." comment: The administered scale (CAPS) is a validated PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: " The 2-way mixed single-measures (consistency) intraclass correlation coefficient for CAPS severity among all assessors over 20 randomly selected cases was 0.81." comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points

4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "Assessors were blinded to treatment allocation. [...] Assessors avoided contact with the therapists and other caregivers. With these procedures, 27 incidents of unblinding occurred (11 in PE, 11 in EMDR, and 5 in WL). In case of unblinding, another assessor repeated the entire measurement." comment: although the assessors were blind to the participants' condition, there was no blind assessment: According to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.
4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	Quote: "Assessors were blinded to treatment allocation. [...] Assessors avoided contact with the therapists and other caregivers."
4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	N	comment: The risk of bias is lowered by the fact that the comparator was also an active intervention. In addition, the active comparator is also a psychotherapeutic intervention and participants received the same number of sessions with the same session duration making them even more comparable. Thus, the probability that participants have answered according to their beliefs/expectations about the intervention effect is regarded low.
Risk of bias judgement	**Low**	In theory, participants might have known their assigned intervention, thus raising the question of bias due to answering interview questions based on their beliefs/expenctations about the intervention effect. This risk of bias, however, is lowered by the fact that the comparator was also an active intervention. In addition, the active comparator was also a psychotherapeutic intervention (common elements) and participants received the

			same number of sessions with the same session duration making both conditions even more comparable. Thus, bias due to participants' knowledge of the intervention is regarded low.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	Quote: "The trial design was approved by the medical ethics committee of the VU University Medical Center and was registered at isrctn.com (ISRCTN79584912)." comment: information and analysis plan from the published study protocol (2013) is consistent with analysis reported in the trial report: de Bont, P. A., van den Berg, D. P., van der Vleugel, B. M., de Roos, C., Mulder, C. L., Becker, E. S., ... & van Minnen, A. (2013). A multi-site single blind clinical study to compare the effects of prolonged exposure, eye movement desensitization and reprocessing and waiting list on patients with a current diagnosis of psychosis and co morbid post traumatic stress disorder: study protocol for the randomized controlled trial Treating Trauma in Psychosis. Trials, 14(1), 151.
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	see comment 5.1
	5.3 ... multiple eligible analyses of the data?	N	see comment 5.1
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Some concerns**	

Unique ID	21	Study ID	220102	Assessor	R
Ref or Label	van den Berg 2015	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	PE	Comparator	WL	Source	Journal article(s) with results of the trial; Trial protocol; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "An independent randomization bureau randomized the treatment condition using stratified randomization blocks per therapist with equal strata sizes.Therapists confirmed the treatment assignment in writing. Data were stored at the study coordination center."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: " At baseline, there were no significant differences between the groups in any of the demographic or clinical characteristics."
	Risk of bias judgement	**Low**	independent randomization bureau randomized the treatment condition using stratified randomization blocks, no significant differences between the groups in any of the demographic or clinical characteristics.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists (and participants) may necessarily have been aware of the assigned condition
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	

			see Figure 1 flow chart comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol that might lead to bias
2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	N		Quote: "After treatment and at the 6-month follow-up, patient files were reviewed to check whether trauma-focused treatments had taken place and if there had been any changes in the prescribed medications, as well as for any deviations from standard care." "Adherence to protocols was rated as good or excellent in 91.2%of PE sessions and 97.1% of EMDR sessions." "There were no differences between groups in additional support provided by caregivers. Groups did not differ in the percentage of participants receiving additional non–trauma focused psychotherapy during treatment (17.0% in PE, 20.8% in EMDR, and 21.3% in WL)"
2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA		
2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA		
2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y		Quote: "Continuous variables were analyzed on an intent-to-treat basis with linear mixed models (LMMs). [...] Analyses of completers and intent-to-treat analyses with last observation carried forward (with missing data on loss of diagnosis conservatively replaced with a negative value [ie, no loss of diagnosis]) were performed to test the robustness of the findings."
2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA		
Risk of bias judgement	**Low**		no blinding of participants/therapists; no deviations; ITT analysis.

Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	see Figure 1 flow-chart comment: post-assessment of 47 out of 53 participants in PE group, and of 39 out of 44 participants in WL group
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	Quote: "Sensitivity Analyses: Completer analyses were performed (n=113), among which no baseline differences were observed between groups in any of the demographic or clinical characteristics. In addition, intent-to-treat analyses with last observation carried forward were performed (n = 155). All results for the CAPS, PSS-SR, and PTCI were similar to the results from the intent-to-treat analyses, thereby underlining the robustness of the findings." comment: the information in the quote on "sensitivity analyses" does not seem sufficient for the judgement that there is evidence that the result was not biased by missing outcome data. no information is provided on reasons for dropout. Missings might not be missing at random, which would require more complex methods for sensitivity analyses (e.g. selection models or pattern mixed models).
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NI	Quote: "There were 2 severe adverse events in PE, 1 in EMDR, and 4 in WL. However, none of the severe adverse events were judged to have been induced by the study."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	comment: the information in the quote on "sensitivity analyses" does not seem sufficient for the judgement that there is evidence that the result was not biased by missing outcome data (3.2). no information is provided on reasons for dropout. Missings might not be missing at random, calling for more complex methods for

			sensitivity analyses (e.g. selection models or pattern mixed models). comment: There was no difference in dropout between the PE (13 participants [24.5%]) and WL (8 participants [17.0%]) group. However, given the fact that reasons for dropout are not reported, there are some concerns
	Risk of bias judgement	**Some concerns**	no difference in dropout; no information on reasons for dropout.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote from the published study protocol: "The CAPS is considered the gold standard to diagnose posttraumatic stress disorder as defined in the DSM-IV-TR and to establish its severity. A review of the empirical literature on psychometric properties of the CAPS [76] indicates that the CAPS has excellent reliability (>0.90), yielding consistent scores across items, raters and testing occasions. There is also strong evidence of validity: the CAPS has excellent (>0.90) convergent and discriminant validity, diagnostic utility, and sensitivity to clinical change."comment: The administered scale (CAPS) is a validated PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: " The 2-way mixed single-measures (consistency) intraclass correlation coefficient for CAPS severity among all assessors over 20 randomly selected cases was 0.81." comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points

	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "Assessors were blinded to treatment allocation. [...] Assessors avoided contact with the therapists and other caregivers. With these procedures, 27 incidents of unblinding occurred (11 in PE, 11 in EMDR, and 5 in WL). In case of unblinding, another assessor repeated the entire measurement." comment: although the assessors were blind to the participants' condition, there was no blind assessment: According to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	NI	comment: the risk of bias due to knowledge of the intervention is rated higher if the comparator is no treatment (here: WL) than when the comparator is another active intervention.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NI	
	Risk of bias judgement	**High**	comment: the risk of bias due to knowledge of the intervention is rated higher if the comparator is no treatment (here: WL) than when the comparator is another active intervention.

Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	Quote: "The trial design was approved by the medical ethics committee of the VU University Medical Center and was registered at isrctn.com (ISRCTN79584912)." comment: information and analysis plan from the published study protocol (2013) is consistent with analysis reported in the trial report: de Bont, P. A., van den Berg, D. P., van der Vleugel, B. M., de Roos, C., Mulder, C. L., Becker, E. S., ... & van Minnen, A. (2013). A multi-site single blind clinical study to compare the effects of prolonged exposure, eye movement desensitization and reprocessing and waiting list on patients with a current diagnosis of psychosis and co morbid post traumatic stress disorder: study protocol for the randomized controlled trial Treating Trauma in Psychosis. Trials, 14(1), 151.
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	see comment 5.1
	5.3 ... multiple eligible analyses of the data?	N	see comment 5.1
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**High**	comment: possible risk of bias due to knowledge of the intervention

Unique ID	22	**Study ID**	220103	**Assessor**	R
Ref or Label	van den Berg 2015	**Aim**	assignment to intervention (the		

			'intention-to-treat' effect)		
Experimental	EMDR	**Comparator**	WL	**Source**	Journal article(s) with results of the trial; Trial protocol; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
Outcome	clinician-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "An independent randomization bureau randomized the treatment condition using stratified randomization blocks per therapist with equal strata sizes.Therapists confirmed the treatment assignment in writing. Data were stored at the study coordination center."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: " At baseline, there were no significant differences between the groups in any of the demographic or clinical characteristics."
	Risk of bias judgement	**Low**	independent randomization bureau randomized the treatment condition using stratified randomization blocks, no significant differences between the groups in any of the demographic or clinical characteristics.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists (and participants) were probably necessarily aware of the assigned condition
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	

			see Figure 1 flow chart
2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	N		comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol that might lead to bias
			Quote: "After treatment and at the 6-month follow-up, patient files were reviewed to check whether trauma-focused treatments had taken place and if there had been any changes in the prescribed medications, as well as for any deviations from standard care."
			"Adherence to protocols was rated as good or excellent in 91.2%of PE sessions and 97.1% of EMDR sessions."
			"There were no differences between groups in additional support provided by caregivers. Groups did not differ in the percentage of participants receiving additional non–trauma focused psychotherapy during treatment (17.0% in PE, 20.8% in EMDR, and 21.3% in WL)"
2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA		
2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA		
2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y		Quote: "Continuous variables were analyzed on an intent-to-treat basis with linear mixed models (LMMs). [...] Analyses of completers and intent-to-treat analyses with last observation carried forward (with missing data on loss of diagnosis conservatively replaced with a negative value [ie, no loss of diagnosis]) were performed to test the robustness of the findings."
2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA		
Risk of bias judgement	**Low**		no blinding of participants/therapists; no deviations; ITT analysis.

Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	see Figure 1 flow-chart comment: post-assessment of 44 out of 55 participants in EMDR group, and of 39 out of 44 participants in the WL group
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	Quote: "Sensitivity Analyses: Completer analyses were performed (n=113), among which no baseline differences were observed between groups in any of the demographic or clinical characteristics. In addition, intent-to-treat analyses with last observation carried forward were performed (n = 155). All results for the CAPS, PSS-SR, and PTCI were similar to the results from the intent-to-treat analyses, thereby underlining the robustness of the findings." comment: the information in the quote on "sensitivity analyses" does not seem sufficient for the judgement that there is evidence that the result was not biased by missing outcome data. no information is provided on reasons for dropout. Missings might not be missing at random, which would require more complex methods for sensitivity analyses (e.g. selection models or pattern mixed models).
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NI	Quote: "There were 2 severe adverse events in PE, 1 in EMDR, and 4 in WL. However, none of the severe adverse events were judged to have been induced by the study."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	comment: the information in the quote on "sensitivity analyses" does not seem sufficient for the judgement that there is evidence that the result was not biased by missing outcome data (3.2). no information is provided on reasons for dropout. Missings might not be missing at random, calling for more complex methods for sensitivity analyses (e.g. selection models or pattern mixed models). comment: There was no difference in dropout between

			the EMDR (11 participants [20.0%]) and WL (8 participants [17.0%]) group. However, given the fact that reasons for dropout are not reported, there are some concerns
	Risk of bias judgement	**Some concerns**	no difference in dropout; no information on reasons for dropout.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote from the published study protocol: "The CAPS is considered the gold standard to diagnose posttraumatic stress disorder as defined in the DSM-IV-TR and to establish its severity. A review of the empirical literature on psychometric properties of the CAPS [76] indicates that the CAPS has excellent reliability (>0.90), yielding consistent scores across items, raters and testing occasions. There is also strong evidence of validity: the CAPS has excellent (>0.90) convergent and discriminant validity, diagnostic utility, and sensitivity to clinical change." comment: The administered scale (CAPS) is a validated PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: " The 2-way mixed single-measures (consistency) intraclass correlation coefficient for CAPS severity among all assessors over 20 randomly selected cases was 0.81." comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points

	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "Assessors were blinded to treatment allocation. [...] Assessors avoided contact with the therapists and other caregivers. With these procedures, 27 incidents of unblinding occurred (11 in PE, 11 in EMDR, and 5 in WL). In case of unblinding, another assessor repeated the entire measurement."

comment: although the assessors were blind to the participants' condition, there was no blind assessment: According to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary. |
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	NI	comment: the risk of bias due to knowledge of the intervention is rated higher if the comparator is no treatment (here: WL) than when the comparator is another active intervention.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NI	
	Risk of bias judgement	**High**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	Quote: "The trial design was approved by the medical ethics committee of the VU University Medical Center and was registered at isrctn.com (ISRCTN79584912)."comment: information and analysis plan from the published study protocol (2013) is consistent with analysis reported in the trial report:de Bont, P. A., van den Berg, D. P., van der Vleugel, B. M., de Roos, C., Mulder, C. L., Becker, E. S., ... & van Minnen, A. (2013). A multi-site single blind clinical study to compare the effects of prolonged exposure, eye movement desensitization and reprocessing and waiting list on patients with a current diagnosis of psychosis and co morbid post traumatic stress disorder: study protocol for the randomized controlled trial Treating Trauma in Psychosis. Trials, 14(1), 151.

	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	see comment 5.1
	5.3 ... multiple eligible analyses of the data?	N	see comment 5.1
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**High**	comment: possible risk of bias due to participants' knowledge of the intervention

Unique ID	26	**Study ID**	260101	**Assessor**	R
Ref or Label	Duffy 2007	**Aim**	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	CT	**Comparator**	WL	**Source**	Journal article(s) with results of the trial; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
Outcome	self-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomizatio n process	1.1 Was the allocation sequence random?	Y	Quote: "An independent office allocated patients to immediate therapy or to wait followed by therapy on a stratified random basis using the minimisation method of Pocock. Assessors were not aware of the allocation algorithm."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "The groups were similar at baseline for personal details, psychiatric status, trauma history, and previous treatments (table1)."
	Risk of bias judgement	**Some concerns**	stratified randomization using the minimisation method, insufficient/ambiguous information regarding allocation concealment; no substantial baseline differences.
	2.1.Were participants aware of their assigned intervention during the trial?	PY	

Bias due to deviations from intended interventions	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	comment: therapists and participants were assumingly necessarily aware of the assigned condition
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	Quote: "The variation in outcome associated with different therapists (14%) was much larger than that observed in a randomised controlled trial of the same treatment (<1%), which was run in a university research clinic with therapists who had received extensive training to protocol." comment: as therapists were not blinded it is possible that therapist effects were caused by therapist allegiance (or differences in training). On the other hand, the fact that the therapist effect caused variation within the CT group, not between the CT and WL group, might argue against a biased effect of assignment to intervention (as adressed here)
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: " We carried out analyses on an intention to treat basis and on patients who completed the study [...]" comment: no further details on method provided
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no major deviations, no blinding, ITT analysis.

Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	see flow-chart Quote: "Twelve patients (21%) dropped out." comment: 12-week post-assessment of 20 out of 29 participants randomized to CT; and of 29 out of 29 randomized to WL
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	comment: no analysis correcting for bias or sensitivity analysis or the like reported; visual inspection of ITT vs. completers data reveals descriptive differences in effect sizes (significance not tested): MD (between groups) between adjusted means on PDS outcome measure: ITT 9.6 (3.6 to 15.6); completers 16.9 (10.9 to 23.0);
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	Quote: " The case notes for the 12 patients who dropped out in the present trial were reviewed. Four were related to the unique context of Northern Ireland (threats to self or family linked to the civil conflict), four were associated with non-adherence to the therapy protocol in relation to imaginal reliving and behavioural experiments (not tackling beliefs, inadequate or inappropriate preparation), two were due to illness or trauma in the family, and two were unknown."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PY	". Some patients seem to have dropped out of the study as a result of problems in running a treatment centre that serves communities where terrorism and other civil conflict related fears and suspicions are still present. However, other patients seem to have dropped out because of suboptimal delivery of the treatment protocol, in particular omitting to use cognitive techniques to tackle some patients' extreme beliefs about the adverse effects of imaginal reliving."

			comment: in addition to reasons for dropout that were related to the treatment condition, the proportions of dropouts between the groups (0 in WL vs. 9 in CT) were unequal, suggesting that missingness in the outcome might depend on its true value
	Risk of bias judgement	**High**	in addition to the fact that reasons for dropout that were related to the treatment, the proportions of dropouts between the groups were unequal.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	comment: The administered scale (PSD) is a validated PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points;
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	comment: there were no independent assessor ratings in this study and the PDS is a self-report questionnaire; thus, the 'assessors' were -in this case-the participants themselves (who were probably aware of their intervention)
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered accoring to their beliefs/expectations about the intervention effect
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NI	
	Risk of bias judgement	**High**	comment: the risk of bias due to knowledge of the intervention is rated higher if the comparator is no treatment (here: WL) than when the comparator is another active intervention
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	Quote: "Trial registration Current Controlled Trials ISRCTN16228473." comment: but the trial registration was done retrospectively, which might raise some concerns
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest

	5.3 ... multiple eligible analyses of the data?	PN	comment: both ITT and completers data is available; generally the risk of bias due to mutiple eligible analyses of the data is low as the results assessed here are raw values (means, SDs); both unadjusted and baseline adjusted effect estimates are reported
	Risk of bias judgement	**Some concerns**	
Overall bias	**Risk of bias judgement**	**High**	there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered accoring to their beliefs/expectations about the intervention effect (risk is especially relevant beuase comparator is a passive control condition). Also, in addition to the fact that reasons for some of the dropout were related to the treatment condition, the proportions of dropouts between the groups were unequal, suggesting that missingness in the outcome could depend on its true value. Last but not least, the trial registration was done retrospectively, not prospectively

Unique ID	37	**Study ID**	290106	**Assessor**	R	
Ref or Label	Ehlers 2004	**Aim**	assignment to intervention (the 'intention-to-treat' effect)			
Experimental	CT	**Comparator**	WL	**Source**	Journal article(s) with results of the trial	
Outcome	clinician-rated PTSD	**Results**	SMD(between)	**Weight**	1	
Domain	**Signalling question**			**Response**	**Comments**	
	1.1 Was the allocation sequence random?			NI		

Bias arising from the randomization process	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	Quote: "Patients were randomly allocated to either immediate cognitive therapy (CT, N = 14) or a 13-week waitlist (WL, N = 14) condition." comment: no further information on radomization or allocation concealment reported
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	Quote: "The groups were comparable in demographic and trauma characteristics (all p's>.26). As shown in Table 3, self-reported symptom severity was also similar on all measures (all p's>.17). However, independent assessors rated the CT group as more severe on the Clinician-Administered PTSD Scale" comment: although on one measure groups differed at baseline, there is no substantial excess in statistically significant differences in baseline characteristics between intervention groups, beyond that expected by chance
	Risk of bias judgement	**Some concerns**	insufficient information on radomization or allocation concealment; no substantial baseline differences betw. groups.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists (and participants) were probably necessarily aware of the assigned condition
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	

	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	comment: all participants received the treatment they were assigned to; no dropouts; all participants randomized were included in the analysis and post-assessment data was available for all participants
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding of participants/therapists; no deviations, ITT analysis.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	Quote: "No patient dropped out."
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	no missing data
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: " A random sample of 38 CAPS interviews (eight different interviewers) from the present and a related study (Ehlers et al., 2003) was rated by a second clinician (seven different raters). Results indicated very good reliability for the PTSD diagnosis, kappa=.94, and total severity score, r = .96." comment: : The administered scale (CAPS) is a validated PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points

	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	Quote: "Independent assessors who were not aware of the treatment condition (trained psychologists) gave the CAPS-SX." comment: although the assessors were blind to the participants' condition, there was no blind assessment: According to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is rated higher if the comparator is no treatment (here: WL) than when the comparator is another active intervention. In the report there is no information indicating that the assessor was an independent researcher not involved in the study, which would otherwise lower the risk of bias.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: raw values (means, SDs) (=unadjusted effect size estimates) are reported for all participants randomized. Baseline adjusted results were also reported
	Risk of bias judgement	**Some concerns**	
Overall bias	**Risk of bias judgement**	**High**	comment: there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect (risk is especially relevant because comparator is a passive control condition).

Unique ID	41	Study ID	300102	Assessor	R	
Ref or Label	Ehlers 2014	Aim	assignment to intervention (the 'intention-to-treat' effect)			
Experimental	CT	Comparator	WL	Source	Journal article(s) with results of the trial	
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1	

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "The participants were then randomly allocated to one of the four trial conditions by an independent researcher who was not involved in assessing patients using the minimization procedure (15) to stratify for sex and severity of PTSD symptoms."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "Table 1 summarizes the details on trauma and the clinical, demographic, and treatment characteristics. No group differences were observed in any of the variables."
	Risk of bias judgement	**Some concerns**	randomization using a minimization procedure; no information on allocation concealment; no group differences at baseline.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	Y	Quote: "Participants were not blind to the nature of the treatment"
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	comment: therapists were probably necessarily aware of the assigned condition
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	

	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "All analyses were intention-to-treat using all 121 randomly assigned participants." "Data were collected from all participants, including dropouts."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding of participants/therapists; no deviations, ITT analysis.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	see Figure 1 flow-chart comment: CT: n = 31 randomized, 1 dropout, n= 31 pre- and post-assessed WL: n= 30 randomized, no dropouts, n=30 pre- and post-assessed
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	no missing data
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The CAPS assesses the frequency and severity of each of the PTSD symptoms specified in DSM-IV. Interrater reliability for a PTSD diagnosis was kappa=0.95, and r=0.98 for the total severity score (37 interviews, 14 interviewers, and 14 raters)." comment: The administered scale (CAPS) is a validated PTSD measure and likely to be sensitive to intervention effects

4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
4.3 Were outcome assessors aware of the intervention received by study participants?	Y	Quote: "Participants were not blind to the nature of the treatment" "Independent assessors (trained psychologists) interviewed patients with the Clinician Administered PTSD Scale (CAPS)" comment: from the information reported ["independent assessors"] it is not clear whether assessors were blinded. even if the assessors were blind to the participants' condition, there was no blind assessment: According to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.
4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: In the report there is no information indicating that the assessor was an independent researcher not involved in the study, which would otherwise lower the risk of bias. : the risk of bias due to knowledge of the intervention is rated higher if the comparator is no treatment (here: WL) than when the comparator is another active intervention. In addition, for subjective outcomes such as 'clinical impression of improvement', knowledge of the intervention received could be highly influential.
4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
Risk of bias judgement	**High**	no information on assessor blinding; no information indicating that the assessor was an independent researcher not involved in the study, which would otherwise lower the risk of bias. the comparator is no treatment, again increasing risk of bias. In addition, for subjective outcomes such as 'clinical impression of improvement', knowledge of the intervention received could be highly influential.

Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	Quote: "The trial was registered as ISRCTN 48524925." comment: however, the registration was done retrospectively, which might raise concerns
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	
	5.3 ... multiple eligible analyses of the data?	PN	comment: ITT data for all participants randomized are available; generally, the risk of bias due to mutiple eligible analyses of the data is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Some concerns**	
Overall bias	**Risk of bias judgement**	**High**	high risk of bias due to knowledge of the intervention (at least of participants; possibly also of assessors)

Unique ID	45	**Study ID**	360101	**Assessor**	R
Ref or Label	Foa 2018	**Aim**	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	PE	**Comparator**	PCT	**Source**	Journal article(s) with results of the trial; Trial protocol; Statistical analysis plan (SAP); Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
Outcome	clinician-rated PTSD	**Results**	SMD(between)	**Weight**	1
Domain	**Signalling question**		**Response**		**Comments**
	1.1 Was the allocation sequence random?		PY		

Bias arising from the randomization process	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	Quote: "The randomization sequence was entered by a study statistician into a secure, web-based application using SAS version 9.4 (SAS Institute Inc), which was accessed by the project coordinator on enrollment of each participant."
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "There were no significant treatment group differences on baseline variables." comment: see Table 1.
	Risk of bias judgement	**Some concerns**	unclear allocation concealment, no substantial baseline differences between groups.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	Quote: "Randomization was originally planned as 3:11:11:11 for MCC:massed therapy: spaced therapy: PCT. On January 5, 2012, enrollment in MCC was accelerated by changing the ratio to 1:1:1:1 to allow for preliminary massed therapy vs MCC comparison per Department of Defense request. After 40 participants were randomized to receive MCC, randomization to MCC was discontinued on March 19, 2014, and subsequent participants were assigned to receive massed therapy, spaced therapy, or PCT (1:1:1). Sensitivity analyses were performed to determine if the results were affected by the different randomization patterns. Randomization pattern was dummy coded and then added as a moderator to the analyses.There were no significant main effects or interactions involving randomization pattern, suggesting that the results did not differ between randomization patterns."
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	

	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "Linear mixed models and generalized linear mixed models were used to analyze the data, using SPSS version 23 (IBMSPSS). These models are intent-to-treat and calculate results based on available data without imputation of missing data."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding of participants/therapists; no deviations, ITT analysis.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	comment: see Figure 1. Quote: "110 Randomized to receive spaced PE"- "79 Completed posttreatment (wk 8) follow-up", "31 Did not complete posttreatment follow-up"; "110 Randomized to receive PCT"- "88 Completed posttreatment (wk 8) follow-up", "22 Did not complete posttreatment follow-up"

| 3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data? | PN | Quote: "Pattern mixture modeling was used to evaluate the effects of missing data.This approach compares growth curve parameters of participants with complete vs missing data." "Participants with missing data did not significantly differ from completers on baseline variables. Further, pattern mixture modeling found no significant differences in the change in outcome over time between participants with missing data and those without missing data.""During treatment, dropout was n=0 for MCC, n=15 (13.6%) for massed prolonged exposure therapy, n=27 (24.8%) for spaced prolonged exposure therapy, and n=13 (12.1%) for PCT (no significant differences among active treatments)." "A total of 59 participants in the spaced therapy group (54.1%) reported adverse events (115 events) during treatment." [23 events were study-related] " 13 Requested to drop out [from PE]""Fifty of the 107 participants in the PCT group (46.7%) reported 96 adverse events during treatment; 1 of these adverse events was study-related." "3 Requested to drop out [from PCT]"comment: although overall dropouts did not significantly differ between active treatments, examining those who requested to drop out (not those who dropped out for reasons unrelated to the treatment), there were more such cases in the PE group (n=13) than in the PCT group (n=3). The exact reasons for the "request to drop out" are not reported. A dependency of missingsness in the outcome on its true value (study related adverse events were also more in the PE (study-related AEs=23) group compared to PCT (= 1)) can therefore not be ruled out. A sensitivity analysis was conducted to assess the potential impact of missing outcome data. It can be assumed that the missingness mechanism is non-ignorable (MNAR), so conducting sensitivity analyses based on pattern mixture modelling seems appropriate. However, the risk of bias is rated as unclear ("some concerns") for the following reasons: (1) arguably, in the presence of non-random dropout, a wholly satisfactory analysis of the data is not feasible. (2) no details on their sensitivity analyses are reported. It is unclear what models |

			have been used, or what assumptions were made and whether reasons for dropout were (a) known to the investigators and (b) addressed in sensitivity analysis.
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NI	

	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	Sensitivity analyses were conducted assuming a MNAR mechanism which seems appropriate (which is why the risk of bias is not rated "high"). Because no details are reported regarding these analyses, however, there are concerns.
	Risk of bias judgement	**Some concerns**	Sensitivity analyses were conducted assuming a MNAR mechanism which seems appropriate (which is why the risk of bias is not rated "high"); however, no details on analyses reported, e.g. whether reasons for dropout were adressed as a potential source of bias; hence, some concerns remain.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The PTSD Symptom Scale –Interview (PSS-I) is a 17-item clinical interview that evaluates DSM-IV PTSD symptom frequency and severity and can provide DSM-IV-diagnosis [...].Test-retest reliability (0.80) and interrater reliability (κ=0.91) are excellent. In the current sample, internal consistency averaged α=.79."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "PTSD symptom severity was assessed by independent evaluators blinded to treatment condition, before and after treatment" comment: although the assessors were blind to the participants' condition, there was no blind assessment: According to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: interviewers were blinded; nonetheless, no blind assessment since participants might have been aware of their

	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN	allocation status and might have answered interview questions according to their beliefs/expectations regarding the treatment effect; also, it was not systematically assessed whether blinding of interviewers was successful. However, the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention.
	Risk of bias judgement	**Some concerns**	there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition. Interviewers were blinded. The risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	Quote: "TRIAL REGISTRATION clinicaltrials.gov Identifier: NCT01049516" "The study protocol and statistical analysis plan are included in the Supplement." comment: no concerns about selective reporting after careful inspection of the trial registry record, the study protocol and SAP
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**Some concerns**	no information on allocation concealment. there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition. However, the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention. There are some concerns regarding bias due to missing outcome data: Sensitivity analyses were conducted but because no details about the analyses are reported, there are concerns.

Unique ID	46	Study ID	360102	Assessor	R
Ref or Label	Foa 2018	**Aim**	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	PE	**Comparator**	WL	**Source**	Journal article(s) with results of the trial; Trial protocol; Statistical analysis plan (SAP); Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
Outcome	clinician-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "The randomization sequence was entered by a study statistician into a secure, web-based application using SAS version 9.4 (SAS Institute Inc), which was accessed by the project coordinator on enrollment of each participant."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "There were no significant treatment group differences on baseline variables." comment: see Table 1.
	Risk of bias judgement	Some concerns	unclear allocation concealment
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	

	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	Quote: "Randomization was originally planned as 3:11:11:11 for MCC:massed therapy: spaced therapy: PCT. On January 5, 2012, enrollment in MCC was accelerated by changing the ratio to 1:1:1:1 to allow for preliminary massed therapy vs MCC comparison per Department of Defense request. After 40 participants were randomized to receive MCC, randomization to MCC was discontinued on March 19, 2014, and subsequent participants were assigned to receive massed therapy, spaced therapy, or PCT (1:1:1). Sensitivity analyses were performed to determine if the results were affected by the different randomization patterns. Randomization pattern was dummy coded and then added as a moderator to the analyses.There were no significant main effects or interactions involving randomization pattern, suggesting that the results did not differ between randomization patterns."
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "Linear mixed models and generalized linear mixed models were used to analyze the data, using SPSS version 23 (IBMSPSS). These models are intent-to-treat and calculate results based on available data without imputation of missing data."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding of participants/therapists; no deviations, ITT analysis.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	comment: see Figure 1. Quote: "110 Randomized to receive spaced PE"- "79 Completed posttreatment (wk 8) follow-up", "31 Did not complete posttreatment follow-up";

			"40 Assigned to receive minimal contact"; "40 Completed postassessment (wk 2)"
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	Quote: "Pattern mixture modeling was used to evaluate the effects of missing data and to provide a sensitivity analysis for the effect of missing data.This approach compares growth curve parameters of participants with complete vs missing data." "Participants with missing data did not significantly differ from completers on baseline variables. Further, pattern mixture modeling found no significant differences in the change in outcome over time between participants with missing data and those without missing data." "During treatment, dropout was n=0 for MCC, n=15 (13.6%) for massed prolonged exposure therapy, n=27 (24.8%) for spaced prolonged exposure therapy, and n=13 (12.1%) for PCT (no significant differences among active treatments)." comment: number (proportions) of dropouts differed between PE (n= 31 missings [28.2%]) and WL (n= 0 missings, [0%]). examining those who requested to drop out from PE (ignoring those who dropped out for reasons unrelated to the treatment), there were n=13 cases in the PE group.The exact reasons for the "request to drop out" are not reported. A dependency of missingsness in the outcome on its true value (study related adverse events were also more in the PE (study-related AEs=23) group compared to WL (= 1)) can therefore not be ruled out. A sensitivity analysis was conducted to assess the potential impact of missing outcome data. It can be assumed that the missingness mechanism is non-ignorable (MNAR), so conducting sensitivity analyses based on pattern mixture modelling seems appropriate. However, the risk of bias is rated as unclear ("some concerns") for the following reasons: (1) arguably, in the presence of non-

			random dropout, a wholly satisfactory analysis of the data is not feasible. (2) no details on their sensitivity analyses are reported. It is unclear what models have been used, or what assumptions were made and whether reasons for dropout were (a) known to the investigators and (b) appropriately accounted for in sensitivity analysis.
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	comment: number (proportions) of dropouts differed between PE (n= 31 missings [28.2%]) and WL (n= 0 missings, [0%]). The exact reasons for the "request to drop out" are not reported. A dependency of missingsness in the outcome on its true value (study related adverse events were also more in the PE (study-related AEs=23) group compared to WL (= 1)) can therefore not be ruled out. Sensitivity analyses were conducted assuming a MNAR mechanism which seems appropriate (which is why the risk of bias is not rated "high"). Because no details are reported
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	

			regarding these analyses, however, the risk of bias is rated "unclear".
	Risk of bias judgement	**Some concerns**	Sensitivity analyses were conducted assuming a MNAR mechanism which seems appropriate (which is why the risk of bias is not rated "high"). Because no details are reported regarding these analyses, however, there are concerns.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The PTSD Symptom Scale –Interview (PSS-I) is a 17-item clinical interview that evaluates DSM-IV PTSD symptom frequency and severity and can provide DSM-IV-diagnosis [...].Test-retest reliability (0.80) and interrater reliability (κ=0.91) are excellent. In the current sample, internal consistency averaged α=.79."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "PTSD symptom severity was assessed by independent evaluators blinded to treatment condition, before and after treatment" comment: although the assessors were blind to the participants' condition, there was no blind assessment: According to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is rated higher if the comparator is no treatment (here: WL) than when the comparator is another active intervention. assessment of the outcome could have been influenced because participants were not blind to their condition and might answer according to their beliefs/expectations about the intervention effect
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	

	Risk of bias judgement	High	there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition. the risk of bias due to knowledge of the intervention is high, as the comparator is a passive control condition (here: WL).
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	Quote: "TRIAL REGISTRATION clinicaltrials.gov Identifier: NCT01049516" "The study protocol and statistical analysis plan are included in the Supplement." comment: no concerns about selective reporting after careful inspection of the trial registry, the study protocol and SAP
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	Low	
Overall bias	Risk of bias judgement	High	no information on allocation concealment; risk of bias due to knowledge of intervention. There are some concerns regarding bias due to missing outcome data: Sensitivity analyses were conducted but because no details about the analyses are reported, there are concerns.

Unique ID	47	**Study ID**	360103	**Assessor**	R
Ref or Label	Foa 2018	**Aim**	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	PCT	**Comparator**	WL	**Source**	Journal article(s) with results of the trial; Trial protocol; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)

Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "The randomization sequence was entered by a study statistician into a secure, web-based application using SAS version 9.4 (SAS Institute Inc), which was accessed by the project coordinator on enrollment of each participant."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "There were no significant treatment group differences on baseline variables." comment: see Table 1.
	Risk of bias judgement	**Some concerns**	unclear allocation concealment
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	Quote: "Randomization was originally planned as 3:11:11:11 for MCC:massed therapy: spaced therapy: PCT. On January 5, 2012, enrollment in MCC was accelerated by changing the ratio to 1:1:1:1 to allow for preliminary massed therapy vs MCC comparison per Department of Defense request. After 40 participants were randomized to receive MCC, randomization to MCC was discontinued on March 19, 2014, and subsequent participants were assigned to receive massed therapy, spaced therapy, or PCT (1:1:1). Sensitivity analyses were performed to determine if the results were affected by the different randomization patterns. Randomization pattern was dummy coded and then added as a moderator to the analyses.There were no significant main effects or interactions involving randomization

			pattern, suggesting that the results did not differ between randomization patterns."
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "Pattern mixture modeling was used to evaluate theeffects of missing data and to provide a sensitivity analysis for the effect of missing data. This approach compares growth curve parameters of participants with complete vs missing data. Linear mixed models and generalized linear mixed models were used to analyze the data, using SPSS version 23 (IBMSPSS). These models are intent-to-treat and calculate results based on available data without imputation of missing data."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding of participants/therapists; no deviations, ITT analysis.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	comment: see Figure 1. Quote: "110 Randomized to receive spaced PE"- "79 Completed posttreatment (wk 8) follow-up", "31 Did not complete posttreatment follow-up";

			"40 Assigned to receive minimal contact"; "40 Completed postassessment (wk 2)"
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	Quote: "Pattern mixture modeling was used to evaluate the effects of missing data and to provide a sensitivity analysis for the effect of missing data.This approach compares growth curve parameters of participants with complete vs missing data." "Participants with missing data did not significantly differ from completers on baseline variables. Further, pattern mixture modeling found no significant differences in the change in outcome over time between participants with missing data and those without missing data." "During treatment, dropout was n=0 for MCC, n=15 (13.6%) for massed prolonged exposure therapy, n=27 (24.8%) for spaced prolonged exposure therapy, and n=13 (12.1%) for PCT (no significant differences among active treatments)." comment: number (proportions) of dropouts differed between PCT (n= 22 missings [20.0%]) and WL (n= 0 missings, [0%]). there were n=3 cases in the PCT group who requested to drop out from treatment, n=6 where the relation of dropout with treatment is unclear, and n= 4 who dropped out for reasons that can be assumed to be unrelated to the treatment). The exact reasons for the "request to drop out" are not reported. Reasons for not being post-assessed were "withdrew" (n=14), "no show or canceled" (n=7), and "lost contact" (n=1). A dependency of missingsness in the outcome on its true value can therefore not be ruled out. The number of study related adverse events, however, was equal in both groups (PCT =1; WL = 1). A sensitivity analysis was conducted to assess the potential impact of missing outcome data. It can be assumed that the missingness mechanism is non-ignorable (MNAR), so conducting sensitivity analyses based on

			pattern mixture modelling seems appropriate. However, the risk of bias is rated as unclear ("some concerns") for the following reasons: (1) arguably, in the presence of non-random dropout, a wholly satisfactory analysis of the data is not feasible. (2) no details on their sensitivity analyses are reported.
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	comment: number (proportions) of dropouts differed between PCT (n= 22 missings [20.0%]) and WL (n= 0 missings, [0%]). there
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	were n=3 cases in the PCT group who requested to drop out from treatment, n=6 where the relation of dropout with treatment is unclear, and n= 4 who dropped out for reasons that can be assumed to be unrelated to the treatment). The exact reasons for the "request to drop out" are not reported. Reasons for not being post-assessed were "withdrew" (n=14), "no show or canceled" (n=7), and "lost contact" (n=1). A dependency of missingsness in

			the outcome on its true value can therefore not be ruled out. The number of study related adverse events, however, was equal in both groups (PCT =1; WL = 1).
	Risk of bias judgement	**Some concerns**	unequal proportions of mising outcome data between both groups; reasons fro dropout may be study-related in some cases; Participants with missing data did not significantly differ from completers on baseline variables; authors conducted sensitivity analysis but some concerns remain.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The PTSD Symptom Scale –Interview (PSS-I) is a 17-item clinical interview that evaluates DSM-IV PTSD symptom frequency and severity and can provide DSM-IV-diagnosis [...].Test-retest reliability (0.80) and interrater reliability (κ=0.91) are excellent. In the current sample, internal consistency averaged α=.79."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "PTSD symptom severity was assessed by independent evaluators blinded to treatment condition, before and after treatment" comment: although the assessors were blind to the participants' condition, there was no blind assessment: According to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is rated higher if the comparator is no treatment (here: WL) than when the comparator is another active intervention. assessment of the outcome could have been influenced because participants
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	

			were not blind to their condition and might answer according to their beliefs/expectations about the intervention effect
	Risk of bias judgement	**High**	there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition. the risk of bias due to knowledge of the intervention is high, as the comparator is a passive control condition (here: WL).
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	Quote: "TRIAL REGISTRATION clinicaltrials.gov Identifier: NCT01049516" "The study protocol and statistical analysis plan are included in the Supplement." comment: no concerns about selective reporting after careful inspection of the trial registry, the study protocol and SAP
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	
	5.3 ... multiple eligible analyses of the data?	N	
	Risk of bias judgement	**Low**	
Overall bias	**Risk of bias judgement**	**High**	no information on allocation concealment; risk of bias due to knowledge of intervention; unclear risk of bias due to missing outcome data: Sensitivity analyses were conducted but because no details about the analyses are reported, there are concerns.

Unique ID	51	**Study ID**	350101	**Assessor**	R
Ref or Label	Foa 2005	**Aim**	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	PE	**Comparator**	WL	**Source**	Journal article(s) with results of the trial

Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	PY	Quote: "The study statistician assigned participants who provided informed consent to one of the three conditions using a weighted randomization procedure such that participants were assigned to one of the active treatment conditions at a greater rate than to WL."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	NI	Quote: "We first examined possible pretreatment differences on PSS–I, BDI, SAS—Work (SAS–W), and SAS—Social (SAS–S) scores across treatment groups and sites using a series of separate single factor analyses of variance (ANOVAs). No significant differences between sites emerged. [...] WL (M = 33.3, SD = 6.2) did not differ from PE/CR, t(98) = 1.3, ns, or PE,t(103) 1.0, ns."
			comment: differences in Demographic characteristics were only analysed regarding the two sites. Demographic characteristics are not reported accoring to the treatment conditions. Only analyses of pre-treatment scores are reported. These do not indicate any problem with the randomization process
	Risk of bias judgement	**Some concerns**	no information on allocation concealment; insufficient information regarding baseline differences between intervention groups
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	Y	Quote: "Therapists made contact with the participants and arranged initial therapy appointments with those assigned to active treatment, and they also informed them of the specific treatment condition at the first session. WL participants were informed by phone that they had been assigned to the WL condition."
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias

	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	PY	Quote: ". Of the 210 eligible women who signed consent, 20 withdrew before being assigned a treatment condition, and 11 were removed from the study after randomization. Thus, our intent-to-treat sample consisted of 179 women who signed consent, were randomized to a condition, and were not removed by the investigators." "We conducted separate Group (WL vs. PE/CR vs. PE) x Site (CTSA vs. WOAR) x Time (pre- vs. posttreatment) mixed factorial ANOVAs on the PSS–I,1 BDI, SAS–W, and SAS–S scores for the intent-to-treat sample, substituting pretreatment scores for missing posttreatment scores." comment: see Figure 1: 11 participants were excluded after randomization. However, the reported reasons for exclusion indicate that 10 of the participants were not eligible after all. Only 1 participant was excluded due to non-adherence ("discontinued medication"). The analysis can therefore still be considered appropriate.
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding of participants/therapists; no deviations, ITT analysis.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "The overall dropout rate was 32.4% and was lower for WL (3.8%) than PE/CR (40.5%), [...], and PE (34.2%)"

			comment: see Figure 1: dropouts in WL n=1 (of 26); in PE n=27 (of 79)
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis or the like reported
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: "Completers differed from noncompleters on level of education,
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	x^2(4, N = 177) = 11.8, p < .05, being more likely to have a bachelor's degree or higher (34% vs. 12%) and less likely to have not completed high school (8% vs. 17%). There were trends for completers to be older (M = 32.2, SD = 9.7) than noncompleters (M =29.3, SD= 10.0), t(176) = 1.9, p = .064, and to be employed full time (43% vs. 33%) or to be students (22% vs. 14%) rather than unemployed (17% vs. 35%), x^2(4, N = 176) = 7.9, p< .096. There was a trend for completion rates to differ across traumas, x^2 (2, N= 179) =4.6, p < .099, with 63% of survivors of adult rape, 76% of nonsexual assault, and 81% of childhood sexual abuse completing treatment. Notably, comorbidity, exposure to additional trauma, or direct experience of additional interpersonal violence was not associated with dropout status" "Twelve serious adverse events led to termination in the study, six of which are included in the postrandomization removal category in Figure 1 (4 participants reassaulted, 1 developing a life threatening illness, and 1 death). The remaining six serious adverse events were classified as dropouts (4 had severe depression and suicidal ideation that required immediate intervention, 2 of which were hospitalized, and 2 exhibited extreme dissociative symptoms)." comment: information on adverse events only in total, not for each

			treatment group. no reasons for dropout reported. Hence, it is difficult to assess the likeliness of a dependency of missingness on its true value. As the proportions of missing outcoe data differ between both groups, bias is possible
	Risk of bias judgement	**High**	No reasons for dropout reported; proportions of missing outcoe data differ between both groups.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The PTSD Symptom Scale—Interview (PSS–I; Foa, Riggs, Dancu, & Rothbaum, 1993) is a semistructured interview that consists of 17 items corresponding to the Diagnostic and Statistical Manual of Mental Disorders (4th ed.; DSM–IV; American Psychiatric Association, 1994) PTSD symptoms. [...] Interrater reliability for PTSD diagnosis (.91) and overall severity (r .97) are excellent (Foa et al., 1993). Of the audiotaped PSS–I interviews in the current study, 5% were randomly selected for rating by a second evaluator. The interrater reliability was .94"
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: " All evaluations were conducted by trained doctoral or master's level CTSA clinicians who were blind to study condition. The same evaluators conducted assessments for both the CTSA and the WOAR participants." comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	Quote: "All evaluations were conducted by trained doctoral or master's level CTSA clinicians who were blind to study condition." comment: although the assessors were blind to the participants' condition, there was no blind assessment: According to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the

			participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the outcome is not based solely on self-ratings of unblinded participants but reflects the clinical impression of improvement rated by an expert who was an independent researcher. However, the risk of bias due to knowledge of the intervention is rated higher if the comparator is no treatment (here: WL) than when the comparator is another active intervention. participants might answer according to their beliefs/expectations about the intervention effect.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	although the interviewers were blinded, there is still a risk of bias. Because the comparator is a passive control condition the risk is higher that participants might answer questions according to their beliefs/expectations regarding their assigned condition. This risk can not be fully eliminated by assessment by a blinded clinician.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PY	comment: results were reported for all time points of interest, but not for all outcome measures of interest: the self-report version of the PSS-I (PSS-SR) was also assessed but no results are reported for this outcome measure
	5.3 ... multiple eligible analyses of the data?	PN	comment: both ITT and completers data is available; generally the risk of bias due to multiple eligible analyses of the data is low as the results assessed here are raw values (means, SDs)

	Risk of bias judgement	High	no information on pre-specified analysis plan available; results were reported for all time points of interest, but not for all outcome measures of interest: the self-report version of the PSS-I (PSS-SR) was also assessed but no results are reported for this outcome measure (without justification). The risk of bias due to selective reporting is therefore high
Overall bias	Risk of bias judgement	High	no information on allocation concealment; risk of bias due to participants' knowledge of intervention; No reasons for dropout reported; proportions of missing outcome data differ between both groups. no information regarding pre-specified analysis plan

Unique ID	55	Study ID	380101	Assessor	R
Ref or Label	Ford 2011	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	PCT	Comparator	WL	Source	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "One hundred forty six women (ages 18–45; M=30.7, SD=6.9) completed the screening and baseline assessment and then were randomized (by a study assessor using numbers concealed in sealed envelopes previously prepared by a different study staff member using the Excel random number generator) to WL (N=45), TARGET (N=48), or PCT (N=53)."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	PY	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PY	Quote: "Comparison of the experimental conditions with chi square tests for categorical variables and ANOVA for continuous measures identified no demographic differences and few differences on the outcome measures at baseline (see Table 2).

			PCT and TARGET were lower than WL on the PTCI-S, COPE-Blame, and PSI-D"
	Risk of bias judgement	**Some concerns**	random allocation sequence, sealed envelopes used, baseline differences between groups on clinical measures
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "Intent-to-treat analyses consistent with the CONSORT definition were conducted using mixed model regression in order to include all participants in each analysis regardless of missing data (Bryk & Raudenbush, 1992; Singer, 1998). "
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding; ITT analysis
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	see Figure 1 Quote: "45 Assigned to Control Group. 10 Did Not Complete Post-Wait Interview: 9 No Response/Withdrew, 1 MovedOut of State" "53 Assigned to Receive PCT. 18 Did Not Complete Interview: 12

			No Response/Withdrew, 6 Completed Subsequent Study Interviews"
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	Quote: ".Missing data due to drop outs, missed interviews, or incomplete measures were analyzed using the SPSS Missing Value Analysis program and found to be random except for four variables: the BDI in the PCT and WL conditions, the PTCI in the PCT and TARGET conditions, the IPSI in the TARGET condition, and the STAI in the WL condition. For those variables, conditions, and time points, participants reporting more severe problems were more likely to have missing data at later time points; however, these exceptions to random missing data were equally distributed across conditions and thus no statistical adjustment was deemed necessary because they were unlikely to affect between-group analyses." comment: we can only be sure that there is no bias due to missing outcome data when: (1) the outcome is measured in all participants; (2) the number of participants with missing outcome data is sufficiently small that their outcomes could have made no important difference to the estimated effect of intervention; or (3) sensitivity analyses confirm that plausible values of the missing outcome data could make no important difference to the estimated intervention effect. The analyses descibed in the report are not sufficient to eliminate a risk of bias
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	comment: no exact reasons for withdrawal are reported and it is unclear what reasons those participants had that did not respond (PCT n=12; WL n=9). It is also not clear why the 6 patients in PCT who "completed subsequent study interviews" did not attend post-assessments. When missingness in the outcome is related to its
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	

			true value and, additionally, the effect of the experimental intervention differs from that of the comparator intervention (which is the case here), missing outcome data will lead to bias. It is possible that some of the reasons for dropout are related to the true values of the missing outcome data.
	Risk of bias judgement	**High**	comment: unknown reasons for dropout lead to a risk of bias. Analyses conducted by the authors do not address this potential relationship between missingness in the outcome and its true value.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The CAPS (Blake et al., 1995; Weathers, Keane, & Davidson, 2001) is a reliable and validated structured interview for DSM-IV (American Psychiatric Association, 1994) categorical diagnoses for full and partial (i.e., meets Criterion B and Criterion C or D; Schnurr et al., 2000) PTSD."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: " Independent interrater reliability for the CAPS total score (intraclass correlation=.97 at baseline, .94 at posttest/followup) and detecting full or partial PTSD (92% agreement, κ=.77) was strong, and adequate for distinguishing full versus partial PTSD (82% agreement, κ=.61)." " Interrater reliability was assessed with randomly selected 25% samples of baseline (N=39) and posttest/ follow-up (N=64) interviews by audiotape review by an independent interviewer." comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	Quote: "A failure to ensure that interviewers were blinded to participant assignments at posttest and follow-up assessments made it impossible to determine whether the superiority of PCT and TARGET versus WL on the one structured interview outcome measure, the CAPS, was free from the effects of interviewer expectancies."

			comment: no blinding of interviewers. no blinding of participants
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	comment: the risk of bias due to knowledge of the intervention is rated higher if the comparator is a no-treatment condition (here: WL) than when the comparator is another active intervention.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	In addition, the interviewer knew the participants' treatment condition - For subjective outcomes such as 'clinical impression of improvement', knowledge of the intervention received could be highly influential.
	Risk of bias judgement	**High**	there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. There is even an accumulated risk of bias as the interviewer was also unblinded. The risk is especially high because the comparator was a no-treatment condition
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: ITT data for all participants randomized available; unadjusted and baseline adjusted results reported; generally, the risk of bias due selection of results based on multiple eligible analyses is low as the results assessed here are raw values (means, SDs)

	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for all outcome measures of interest; for all time points of interest; ITT data for all participants randomized available; unadjusted and baseline adjusted results reported; generally, the risk of bias due selection of results based on multiple eligible analyses is low as the results assessed here are raw values (means, SDs). All in all, difficult to assess due to lack of information; hence, some concerns can not be eliminated
Overall bias	**Risk of bias judgement**	**High**	risk of bias due to participants' and interviewers' knowledge of intervention; possible bias due to missing outcome data as reasons for dropout are not reported in sufficient detail; no information regarding pre-specified analysis plan. differences in baseline PTSD scores between groups might lead to a biased Effect size estimate (SMD(pst-post))

Unique ID	56	**Study ID**	390101	**Assessor**	R	
Ref or Label	Galovski 2013	**Aim**	assignment to intervention (the 'intention-to-treat' effect)			
Experimental	CPT	**Comparator**	WL	**Source**	Journal article(s) with results of the trial	
Outcome	clinician-rated PTSD	**Results**	SMD(between)	**Weight**	1	

Domain	**Signalling question**	**Response**	**Comments**
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "If eligible, participants were randomly assigned in a 1:1 ratio using computer generated simple randomization to MCPT or to SMDT following the pre-treatment assessment."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "Randomization was effective as no outcome variable had different baseline values by treatment group (p > .25)."

		Some concerns	random allocation sequence and no baseline differences; but unclear allocation concealment
	Risk of bias judgement		
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	see Figure 1 comment: no deviations reported; no irregularities raising concerns
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	PY	Quote: " First, MCPT was compared to SMDT in the intent-to-treat (ITT) randomization sample (n = 100) to examine pretreatment to post-treatment change." "Consistent with intention-to-treat principles, treatment drop-outs were invited back for post-treatment and subsequent 3-month follow-up assessments." "An ITT philosophy was used for creation of the primary outcome models."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding; no major deviations; ITT analysis.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	see figure 1 Quote: "Although 100 participants were enrolled, 25 of these contributed only one score that could be used in the models."

	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	Quote: "Sensitivity of the estimates to missing data was examined using a model that assumed the data were missing at random that used all available measurements" comment: (1) no further information reported on sensitivity analyses, except of analyses regarding differences in dropouts; (2) assuming a MAR missingness mechanism does not seem appropriate given the differences in characteristics between participants with and without missing data and given the fact that there is no sufficient information regarding reasons for dropout
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	Quote: "Drop-out percentages were not significantly different across the initial randomization conditions ($\chi 2$ (1; N = 91) = 2.47, p = .116, Cramer's V = .165). However, study drop-outs had significantly higher pretreatment CAPS severity (p = 0.028). There were also trends for study drop-outs to be younger (p = 0.081) and to have lower household income (p = 0.074). No differences were found on depression severity or trauma variables. Fifty percent of the 14 participants who dropped out of active treatment stated major, ongoing psychosocial stressors as the reason for leaving therapy. These stressors included issues such as lack of transportation or childcare, home foreclosure, need to move out of state, and imprisonment. The other half of the drop-outs did not report a reason for terminating early."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	comment: although proportions regarding missing t2 outcome data were not significantly different, there is still a risk of bias as 50% of participants who dropped out did not state their reasons for dropout. It is particularly important to note that dropouts had significantly higher pretreatment CAPS scores. It can be assumed that at least in parts data was not missing at random and that it is

			not unlikely that there is a relationship between missingness in the outcome and its true value.
	Risk of bias judgement	**High**	no appropriate sensitivity analyses; unclear reasons for dropout and significantly higher pretreatment CAPS scores in dropouts indicate that the missingness mechanism is probably non-ignorable
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "Clinician-Administered PTSD Scale: (CAPS; Blake et al., 1990) is a widely used clinician-administered diagnostic instrument designed to assess the frequency and intensity of the 17 PTSD symptoms, as well as clinician-rated validity of client report and symptom severity and improvement. [...] The CAPS has demonstrated excellent reliability and validity (Weathers, Keane, & Davidson, 2001). Internal consistency for the 17 PTSD symptoms in current study was high (α = .93)."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "Inter-rater reliability was conducted for a random sample of interviews in the present study (29 CAPS and 25 SCID). Reliability among coders was high for the CAPS, [κ (current diagnosis) = 1.00; r (total score) = .91]" comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "Post-treatment and follow-up assessments were conducted by raters blind to both randomization and drop-out status" comment: although the assessors were blind to the participants' condition, there was no blind assessment as participants were not blinded and might have answered interview questions accoring to their beliefs regarding the treatment condition they were assigned to

	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	comment: could have been influenced because participants were not blind to their condition. The risk of bias due to knowledge of the intervention is higher, as the comparator is a passive control condition (here: WL). It would be lower if the comparator was another active intervention.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	although the assessors were blind to the participants' condition, there was no blind assessment as participants were not blinded and might have answered interview questions accoring to their beliefs regarding the treatment condition they were assigned to.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: ITT data for all participants randomized available; generally, the risk of bias due selection of results based on multiple eligible analyses is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Some concerns**	
Overall bias	**Risk of bias judgement**	**High**	no information on allocation concealment; risk of bias due to participants' knowledge of intervention; possible bias due to missing outcome data as reasons for dropout are not reported in sufficient detail; no information regarding pre-specified analysis plan

Unique ID	58	Study ID	400101	Assessor	R
Ref or Label	Gerbarg 2013	**Aim**	assignment to intervention (the 'intention-to-treat' effect)		

Experimental	MBI	Comparator	WL	Source	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "Subjects meeting inclusion and exclusion criteria were ranked according to baseline CAPS and assigned to a control or treatment group using a computer generated randomization procedure. Numbers were used to assign one participant from each successive pair into each of the study groups."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	comment baseline characteristics are only presented for participants' basline scores on clinical measures and age. There are no differences. However, it would have been desirable to have more information regading demographic variables
	Risk of bias judgement	Some concerns	random allocation sequence but unclear allocation concealment
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	Y	Quote: "It was not possible to blind the participants about the intervention, but each participant was asked not to disclose his or her assigned randomization group to the CAPS assessor. "
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "A Last Observation carried Forward (LOCF) analysis was undertaken with all participants for whom an assessment on the scales utilized was recorded at baseline."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	

	Risk of bias judgement	Low	no blinding; no major deviations; ITT analysis.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: " Thus, a total of 25 participants completed the study." [out of 31 participants; 80.6%] "When participants did not have a score following their baseline evaluation, their last reported/ observed score was used"
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis or the like reported. Imputing the outcome variable through the 'last-observation-carried-forward' method should not be assumed to correct for bias due to missing outcome data.
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NI	Quote: "Of the 16 assigned to the Intervention Group, one could not perform yoga breathing due to severe dyspnea secondary to advanced chronic obstructive pulmonary disease (COPD). He subsequently withdrew consent from the study. Another subject refused to participate in testing after the intervention, in part, related to specific worries that documentation of his improvements could affect his disability benefits. Out of 15 assigned to the Control Group, a total of four participants withdrew from the study during the waiting period. Reasons included work schedule conflicts, social anxiety, pre-existing severe dyspnea, and failure to participate in the baseline testing."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	comment: the significance of the difference in the proportion of dropouts (MBI 12.5%; WL 26.7%) was not tested or not reported. The reported reasons for dropout indicate that for some of the dropouts reasons may be related to the assigned treatment condition or withdrawal was related to the participant's health status. It is possible that missingness in the outcome depends on its true value but not necesserily likely. There's a lot of uncertainty and it is difficult to evaluate the risk of bias based on the information provided. There are definitely concerns that results might be biased due to missing outcome data

	Risk of bias judgement	**High**	not enough information provided in order to evaluate the likeliness that missingness in the outcome depended on its true value. So on the one hand, results might be biased favouring the experimental. At the same time, in view of the fact that missing data was imputed using the LOCF method, results might as well be biased towards the null. The likeliness as well as the predicted direction of bias remain uncertain.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The Clinician Administered PTSD Scale is a 30-item structured interview corresponding to DSMIV criteria for PTSD. [...] It is considered the "gold standard" in assessing PTSD."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	Quote: "It was not possible to blind the participants about the intervention, but each participant was asked not to disclose his or her assigned randomization group to the CAPS assessor. The CAPS administrator and those who scored tests and performed data entry were blinded to group assignment. The CAPS assessor was not a SKY practitioner or teacher." comment: although the assessors were blind to the participants' condition, there was no blind assessment as participants were not blinded and might have answered interview questions according to their beliefs regarding the treatment condition, they were assigned to.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the outcome is not based solely on self-ratings of unblinded participants but reflects the clinical impression of improvement rated by an expert who was no teacher of that mindfulness-based intervention, therefore lowering the risk of allegiance to this treatment. On the other hand, there is no information indicating that this interviewer was an independent researcher not involved in the study which would lower the risk of
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NI	

			bias. Also, the comparator is a passive control condition increasing the risk that participants might answer questions according to their beliefs/expectations regarding their assigned condition. This risk can not be fully eliminated by assessment by a blinded interviewer, especially not since he or she might potentially have been involved in the study.
	Risk of bias judgement	**High**	there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might answer according to their beliefs/expectations about the intervention effect. This risk can not be fully eliminated by assessment through a blinded interviewer (who might be involved in the study)
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: ITT data for all participants randomized available; generally, the risk of bias due selection of results based on multiple eligible analyses is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for all outcome measures of interest; for all time points of interest; ITT data for all participants randomized available; generally, the risk of bias due selection of results based on multiple eligible analyses is low as the results assessed here are raw values (means, SDs). All in all, difficult to assess due to lack of information; hence, there are concerns
Overall bias	**Risk of bias judgement**	**High**	no information allocation concealment; high risk of bias due to participants' knowledge of intervention; potential bias due to missing outcome data/ not enough information for reliable assessment of the dependency of missingness in the outcome on its true value; no information regarding pre-specified analysis plan

Unique ID	63	Study ID	530101	Assessor	R
Ref or Label	Hijazi 2012	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	NET	Comparator	WL	Source	Non-commercial trial registry record (e.g. ClinicalTrials.gov record); "Grey literature" (e.g. unpublished thesis)
Outcome	self-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "The randomization scheme was constructed by a team member not involved with recruiting or running participants using a computer randomization website (randomization.com)"
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	PY	" Undergraduate research assistants made envelopes that included a randomization number and a slip of paper with the corresponding experimental or control condition as per the original randomization scheme." "The randomization scheme was concealed from the research assistants until they were on the phone with the participant and consulted the randomization scheme [...]"
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "To determine the success of randomization, the experimental and control groups were compared on demographics and baseline measures. The two groups did not differ significantly or meaningfully on any of the demographic measures or most of the baseline measures, but the experimental

			group was more likely to have used English language training services (p < .05), see Table 2, Table 3, and Table 4."
	Risk of bias judgement	**Low**	computer randomization website (randomization.com) used; sealed envelopes used for allocation concealment, no baseline differences
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "Intent to treat analyses were used to compare the effects of the group on the entire sample. For participants who were missing either of the follow-up visits, their values from the last visit were carried forward to replace the missing values. For participants who were missing a scale within a visit, imputation was conducted to replace that value."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	

	Risk of bias judgement	**Low**	Quote: "Furthermore, we did not monitor in either of the groups the nature and frequency of any other relevant services they received, such as medication management or psychological interventions. These other services might have influenced outcomes in unknown ways." comment: Although not assessed when rating the effect of assignment to intervention, it should be noted, that the study results might be biased due to non-protocol interventions
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	Quote: "No participants formally dropped from the study, and all participants completed at least one of the two follow-up assessments: 2-month follow-up questionnaires were received from 96% (n = 51) of the sample" Quote from figure 1 flow-chart: "NET Group n = 35, 2-month Follow-up n = 34"; "Waitlist Control Group n = 18, 2-month Follow-up n = 17" comment: two missing post-assessment values; one missing data point in each of the groups; no dropouts
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	two missing values at post-assessment (one per group), no formal dropouts from the study

Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	PN	Quote: "Exposure to traumatic events and post traumatic stress symptoms was assessed using parts of an Arabic adaptation of the Harvard Trauma Questionnaire (Shoeb, Weinstein & Mollica, 2007). [...] Part D, the PTSD symptoms portion of the scale, was

used in this study as the primary outcome, at baseline and at both follow-up points. In this study, Part D demonstrated excellent reliability at baseline (Cronbach's alpha = .92), 2-month follow up (Cronbach's alpha = .97), and 4-month follow-up (Cronbach's alpha = .97)." |
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	comment: the 'assessors' were -in this case- the participants themselves (who were most likely aware of their intervention).
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	the risk of bias due to knowledge of the intervention is rated higher because the comparator is a no-treatment condition (here: WL) (in contrast to the comparator being another active intervention).
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might answer according to their beliefs/expectations about the intervention effect. The lack of blinding is particularly problematic as the comparator is a passive no-treatment condition.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: "The trial was registered with clinicaltrials.gov (NCT01288690)"

comment: Trial Registration.The researchers' pre-specified intentions are available in sufficient detail to believe that outcome measurements and analyses (of means, SDs) can be compared with those published in the report |

	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: see above; also, , the risk of bias due selection of results based on multiple eligible analyses is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Low**	The trial was prospectively registered at ClinicalTrials.gov. Examination of the history of changes and a comparison of this data with the reported information indicates that, all in all, researchers adhered to their pre-specified intentions.
Overall bias	**Risk of bias judgement**	**High**	sufficient information on randomization procedure and allocation concealment to trust that the risk of bias in this domain is low; risk of bias due to participants' knowledge of intervention; trial was pre-registered on ClinicalTrials.gov

Unique ID	64	Study ID	540101	Assessor	R
Ref or Label	Hijazi 2014	**Aim**	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	NET	**Comparator**	WL	**Source**	Journal article(s) with results of the trial; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
Outcome	self-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: " A few days later, the assistant telephoned the participant and asked if he or she was willing to continue participating in the study. If so, the assistant (heretofore blind to condition assignment) opened a sealed envelope and informed the participant when he or she would be getting the treatment. The computerized scheme was stratified by recruitment site (agency)
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	PY	

			and assistant, and randomized the two conditions in blocks of six in a 2:1 ratio (intervention: control)"
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "The two conditions did not differ significantly on any demographic or baseline outcome measure, suggesting successful randomization."
	Risk of bias judgement	**Low**	random allocation sequence, allocation concealment using sealed envelopes, no baseline differences between groups
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: " Our primary analyses were "intent-to-treat," meaning that we retained all 63 participants, regardless of how many intervention or follow-up assessment sessions they completed. Any missing follow-up data were replaced using the multiple imputation procedure in SPSS 20.0."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	therapists and participants were probably necessarily aware of the assigned condition; ITT analysis using multiple imputation
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	comment: see figure 1 flow-chart. Of the 41 participants randomized to NET, 38 completed post-assessment; of the 22 randomized to WL, 21 completed post-assessment
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	

	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NI	Quote: ""one person (in brief NET) was lost to both follow-ups"
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	"The two conditions did not differ on the percentage of participants missing a follow-up (p = .28), and participants who completed both follow-ups did not differ on demographics or baseline values from participants who missed a follow-up (results not shown)."
			comment: 3 participants in the NET group were not post-assessed at t2, however, only "one person (in brief NET) was lost to both follow-ups", and only two participants did not complete the 3 NET sessions: one became employed (reason not related to the treatment condition), for the other one the reason for dropout is not reported. Proportions of missing data did not differ between the groups. The fact that results (or at least p-values) are not reported for significance tests of differences between participant characteristics (dropouts vs. completers); nevertheless, in the light of the information described above, the small amount and equal proportions of missing data, it is unlikely that missingness in the outcome depended on its true value and that results are biased due to missing outcome data
	Risk of bias judgement	**Some concerns**	in the light of the information described above, the small amount and equal proportions of missing data, it is unlikely that missingness in the outcome depended on its true value/ that results are biased due to missing outcome data
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	PN	Quote: "Posttraumatic stress symptoms were assessed by using two sections of the HTQ, which was previously translated into Arabic and used with Iraqi refugees in the U.S. (Shoeb, Weinstein, & Mollica, 2007). [...] This sample's α were .93, 97, and .97."
			comment: The HTQ is a validated PTSD measure and likely to be sensitive to intervention effects

	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	comment: the version of the HTQ used here is a self-report questionnaire; thus, the 'assessors' were -in this case- the participants themselves (who were probably aware of their intervention)
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	Quote: "Second, although randomization to a wait-list condition controlled for some factors, we cannot rule out the possibility that the positive outcomes of brief NET stemmed from other nonspecific factors or biases, such as simply meeting with a caring person, having the same assistant conduct the screening, baseline assessment, and therapy; or demand characteristics to report benefits on self-report measures"
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	comment: there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might answer according to their beliefs/expectations about the intervention effect. the risk of bias due to knowledge of the intervention is particularly high because the comparator is a no-treatment condition (here: WL)
	Risk of bias judgement	**High**	there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might answer according to their beliefs/expectations about the intervention effect. the risk of bias due to knowledge of the intervention is particularly high because the comparator is a no-treatment condition (here: WL)
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: "The study was approved by the Wayne State University Institutional Review Board and registered with clinicaltrials.gov (NCT01288690);"
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest

	5.3 ... multiple eligible analyses of the data?	PN	comment: ITT data for all participants randomized reported; generally, the risk of bias due selection of results based on multiple eligible analyses is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Low**	The trial was prospectively registered at ClinicalTrials.gov. Examination of the history of changes and a comparison of this data with the reported information indicates that, all in all, researchers adhered to their pre-specified intentions.
Overall bias	**Risk of bias judgement**	**High**	random allocation sequence, allocation concealment using sealed envelopes, no baseline differences between groups; ITT analysis using multiple imputation; there is a risk of bias due to knowledge of the intervention as participants were not blind to their condition; The trial was pre-registered at ClinicalTrials.gov

Unique ID	74	Study ID	760102	Assessor	R	
Ref or Label	Kubany 2003	**Aim**	assignment to intervention (the 'intention-to-treat' effect)			
Experimental	CT	**Comparator**	WL	**Source**	Journal article(s) with results of the trial	
Outcome	clinician-rated PTSD	**Results**	SMD(between)	**Weight**	1	
Domain	**Signalling question**			**Response**	**Comments**	
Bias arising from the	1.1 Was the allocation sequence random?			NI	Quote: "After these assessments, the women were randomly assigned to either an Immediate or a Delayed CTT-BW condition."	
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?			NI		

randomization process	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "There were no significant differences on any of the comparisons, suggesting that random assignment was effective in canceling out error related to relevant measured variables." "there were no significant differences in CAPS scores between participants in the Immediate and Delayed C'IT-BW conditions at the initial pretherapy assessment, F(1,30) < 1."
	Risk of bias judgement	**Some concerns**	no information on random sequence generation or allocation concealment; baseline scores on relevant variables did not differ between groups
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	Quote: "Edward Kubany [first authour] served as therapist for all 37 participants."
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as it is hardly possible to achieve blinding when psychological interventions are implemented
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "[...], we conducted intent-to-treat analyses on the data by evaluating outcomes for all participants, using pretreatment data scores for women who started but did not complete treatment (Kazdin, 1998)."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	therapists and participants were probably aware of the assigned condition; no deviations that arose because of the experimental context; ITT analysis used.

Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "Eighteen of 19 women assigned to the Immediate CTT-BW condition completed CTT-BW. Fourteen of 18 women assigned to the Delayed CTT-BW condition completed CTT-BW. Overall, 86% of the 37 women who started CTT-BW (n = 32) completed treatment." comment: only the 32 completers were included in analysis; no post-assessment of dropouts
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	Quote: "To examine the effects of attrition on outcomes, considering noncompleters as treatment failures, we conducted intent-to-treat analyses on the data by evaluating outcomes for all participants, using pretreatment data scores for women who started but did not complete treatment (Kazdin, 1998). Results presented in Table 4 show that, for both the Immediate and Delayed CTT-BW groups, there were large, statistically significant improvements on all treatment-outcome variables, even when pretherapy data for noncompleters were included in the analyses." comment: no sensitivity analyses showing that results are little changed under a range of plausible assumptions about the relationship between missingness in the outcome and its true value. Imputing the outcome variable through methods such as 'last-observation-carried-forward' should not be assumed to correct for bias due to missing outcome data.
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NI	Quote: "Visual comparisons of the numbers for therapy completers and noncompleters do not reveal any pattern of differencesb etween therapy completers and noncompleters."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	comment: no information on the number of dropouts between t1

			and t2 (only the overall number of dropouts in the delayed treatment group is reported - between t1 and t3, t3= after receiving delayed CT treatment). no significance test of differences between completers and dropouts. no information regarding the reasons for dropout. Not enough information to assess the missingness mechanism
	Risk of bias judgement	High	n=5 dropouts (13.5%) without posttest data; no evidence that the result was not biased; authours state that visual inspection of the scores of completers vs. noncompleters on relevant variables suggests there are no differences, however, no significance test was conducted; no information on the proportion of dropouts or on reasons for dropout.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The CAPS (Blake et al., 1990) is a structured inter- view for assessing the symptoms of PTSD according to criteria in DSM-IV. The CAPS was found to have very good diagnostic efficiency when judged against the Struc- tured Clinical Interview for DSM-111-R (Weathers et al., 1992)."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	N	Quote: "After these assessments, the women were randomly as- signed to either an Immediate or a Delayed CTT-BW condition. Two weeks after completing CTT-BW, women in the Immediate CTT-BW condition received their post- therapy assessment. At the same time (about 6 weeks af- ter their initial assessment), women in the Delayed CTT- BW condition received a second pretherapy assessment." comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: " The assessors were blind to participants' condition assignments." comment: although the assessors were blind to the participants'

			condition, there was no blind assessment, for the participants (answering interview questions) were probably aware of their treatment allocation
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	comment: although the interviewers were blinded, there is still a risk of bias. Because the comparator is a passive control condition the risk is higher that participants might have answered questions according to their beliefs/expectations regarding their assigned condition. This risk can not be fully eliminated by assessment by a blinded clinician.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	risk of bias due to participants' knowledge of their treatment allocation, possibly influencing their answers in the interview. Although interviewers were blinded there is still a risk of bias.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest. Generally, a large variety of measures was used and reported
	5.3 ... multiple eligible analyses of the data?	NI	comment: both results from PP analysis and ITT analysis are reported
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure of interest; for all time points of interest; both PP and ITT results are reported. Beyond that, it is difficult to assess the risk of bias in this domain due to lack of information; therefore, some concerns remain.
Overall bias	**Risk of bias judgement**	**High**	no information on randomization procedure and allocation concealment; ITT analysis used; risk of bias due to participants' knowledge of intervention; potential bias due to missing outcome data as number and reasons for dropout are not reported in sufficient detail; little information available to assess risk of bias due to selective reporting.

Unique ID	78	Study ID	770104	Assessor	R	
Ref or Label	Kubany 2004	**Aim**	assignment to intervention (the 'intention-to-treat' effect)			
Experimental	CT	**Comparator**	WL	**Source**	Journal article(s) with results of the trial	
Outcome	clinician-rated PTSD	**Results**	SMD(between)	**Weight**	1	

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	NI	Quote: "Every 2 consecutive women determined to be eligible were randomly assigned either to an immediate CTT-BW condition or to a delayed CTT-BW condition."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "Comparisons, using analyses of variance (ANOVAs) or chisquare tests, were made between the initial scores of participants in the immediate CTT-BW condition and the delayed CTT-BW condition on (a) all the major outcome variables, (b) age, (c) education, (d) ethnicity (White/ethnic minority), (e) medication use (yes/no), (f) concomitant other therapy (yes/no), and (g) number of types of traumatic events reported. There were no significant differences on any of the comparisons, suggesting that random assignment was effective in canceling out error related to relevant measured variables."
	Risk of bias judgement	Some concerns	no information on random sequence generation or allocation concealment; no baseline differences between groups
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as it is hardly possible to achieve blinding when psychological interventions are implemented
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	

	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: there were changes from assigned intervention that are inconsistent with the trial protocol, such as non-adherence (early treatment termination), but these are consistent with what could occur outside the trial
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: " To examine the effects of attrition on outcomes, considering therapy-nonstarters and noncompleters as treatment failures, we conducted intent-totreat analyses on the data by evaluating outcomes for all participants who were randomly assigned, using pretreatment data scores for posttreatment scores for nonstarters and noncompleters"
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	therapists and participants were probably necessarily aware of the assigned intervention; deviations are consistent with what could occur outside the trial context; ITT analysis (LOCF method) used
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "However, posttreatment assessment data were only available for 84 participants" comment: see table 3. 125 participants were randomized. according to table 3, t2 data was available for 85 [not 84, as mentioned in the quote] participants out of 125 randomized, meaning that t2 data was missing for approximately one third of participants.

	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	Quote: " To examine the effects of attrition on outcomes, considering therapy-nonstarters and noncompleters as treatment failures, we conducted intent-totreat analyses on the data by evaluating outcomes for all participants who were randomly assigned, using pretreatment data scores for posttreatment scores for nonstarters and noncompleters (Kazdin, 1994)." comment: no sensitivity analyses showing that results are little changed under a range of plausible assumptions about the relationship between missingness in the outcome and its true value. Imputing the outcome variable through methods such as 'last-observation-carried-forward' should not be assumed to correct for bias due to missing outcome data.
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	Quote: " Compared with completers, women who did not complete CTT-BW were on average younger, less educated, more depressed, more shame prone, and had lower self-esteem at the initial assessment. There were no significant differences between completers and noncompleters in terms of the number of women who were on medication or receiving other therapy"
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	comment: missing T2 data in CT group = 28.57%; in WL group = 35.48%; no significance test was conducted by the authours but visual inspection suggests that proportions of dropouts are approximately equal. The information reported and quoted above indicates that dropouts were more burdened with symptoms (e.g. more shame, more depressed, lower self-esteem) which increases the risk of bias. In addition, reasons for dropout are not reported. It is possible that missingness is related to symptom severity, experiencing adverse events, disappointment in being assigned to the waitlist condition or other reasons related to the treatment. The risk of bias is therefore high

	Risk of bias judgement	**High**	posttest data missing for approximately one third of participants. no evidence that the result was not biased. proportions of dropouts between groups are approximately equal. However, since the overall amount of missing data is considerable, and the fact that there were significant differences between participants with and without missing data raises strong concerns. In addition, reasons for dropout are not reported. The risk of bias is therefore considered high.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: " The CAPS (Blake et al., 1990) is a structured interview for assessing the symptoms of PTSD according to criteria in the Diagnostic and Statistical Manual of Mental Disorders (4th ed.; DSM–IV; American Psychiatric Association, 1994). The CAPS was found to have very good diagnostic efficiency when judged against the Structured Clinical Interview for DSM–III–R (Weathers et al., 1992)." comment: the CAPS is a validated, gold standard scale and likely to be sensitive to treatment effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: " Two weeks after completing CTT-BW, women in the immediate CTT-BW condition received their posttherapy assessment. At the same time (about 6 weeks after their initial assessment), women in the delayed CTT-BW group received a second pretherapy assessment" comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points

	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: " The assessors were blind to participants' condition assignments, and none served as therapists in the study." comment: although the assessors were blind to the participants' condition, there was no blind assessment, for the participants (answering interview questions) were probably aware of their treatment allocation
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: although the interviewers were blinded, there is still a risk of bias. Because the comparator is a passive control condition the risk is higher that participants might have answered questions according to their beliefs/expectations regarding their assigned condition. This risk can not be fully eliminated by assessment by a blinded clinician.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. It is particularly high as the comparator was a no-treatment condition (WL)
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	comment: no pre-specified analysis plan provided
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: detailed rationale for the choice of instruments and measures reported; results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: both ITT and completers data are available; generally, the risk of bias due to multiple eligible analyses of the data is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for all outcome measures of interest; for all time points of interest; ITT data for all participants randomized available; completers data reported; generally, the risk of bias due selection of results based on multiple eligible analyses is low as the results

			assessed here are raw values (means, SDs). All in all, difficult to assess due to lack of information; hence, there are some concerns
Overall bias	**Risk of bias judgement**	**High**	no information on random sequence generation or allocation concealment; no baseline differences between groups but PTSD scores at baseline higher in the WL group which might lead to a biased effect size estimate. ITT analysis used. considerable amount of missing data, significant differences between participants with/without posttest data; no information on reasons for dropout. risk of bias due to participants' knowledge of the intervention. no information regarding pre-specified analysis plan

Unique ID	79	Study ID	870101	Assessor	R
Ref or Label	Markowitz 2015	**Aim**	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	PE	**Comparator**	IPT	**Source**	Journal article(s) with results of the trial; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
Outcome	clinician-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "Eligible individuals who provided written informed consent for the treatment study were randomly assigned to receive prolonged exposure, IPT, or relaxation therapy, in a 4:4:3 ratio. Randomization followed a computer-generated program designed by the study's statistician, who had no patient contact. Randomization was stratified by presence of major depressive disorder (diagnosed according to the SCID, along with a score
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	

			$20 on the 24-item Hamilton Depression Rating Scale [HAM-D] [33]) and implemented in blocks of random sizes (11 or 22)."
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	comment: see table 2; table 3
	Risk of bias judgement	**Some concerns**	computer-generated random sequence; no information on allocation concealment; inspection of baseline characteristics gives no rise to concerns
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	Quote: "[...] and patients were reminded not to identify their therapy or therapist during evaluation."
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as it is hardly possible to achieve blinding when psychological interventions are implemented
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "Analyses followed the intention-to-treat principle. [...] Efficacy of the three treatments with respect to symptom severity was estimated based on longitudinal mixed-effects models (42) using multiple imputation for the missing values"
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	deviations (e.g. non-adherence by participants who did not attend all sessions) did not arise because of the trial context; ITT analysis with multiple imputation used

Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "Attrition was 15% in the IPT group, 29% in the prolonged exposure group, and 34% in the relaxation therapy group (n.s.). Two patients from each treatment condition withdrew after randomization but before beginning therapy."
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	Quote: " We compared participants with missing postrandomization data to those without missing data with respect to baseline characteristics. No comparisons between subjects with and without postrandomization assessment overall, or within treatment groups, were significantly different, and no differences approached clinically meaningful magnitude." "For each variable (score on the CAPS, the Posttraumatic StressScale–Self Report, the HAM-D, the Social Adjustment Scale–Self-Report, the quality of life measure, and the Inventory of Interpersonal Problems), we used the Markov chain Monte Carlo technique to obtain a monotone missing data pattern. We then applied a predictive mean-matching regression method separately for the three treatment groups (44); to increase the likelihood that the missing-at-random assumption is valid, in addition to the previous values of the variable being imputed, we used all other symptom variables and baseline major depression status as predictors in predictive mean-matching regression. Fifty imputed data sets were generated." " The omnibus test assessing whether dropout depended on the interaction between depression status and treatment showed a p value of 0.15. Half of patients who had comorbid depression and were assigned to receive prolonged exposure dropped out: the odds ratio of prolonged exposure attrition with (50%) and without (5.6%) major depression was 17:1 (Table 4). Dropout among depressed patients in the prolonged exposure group tended to be higher than among depressed patients in the IPT group (p=0.086) and higher than dropout among nondepressed patients in the prolonged exposure group (p=0.006)."

			comment: although the methods that were used by the authours to handle missing data consider several potential sources of bias, they do not offer definite evidence that the results was not biased by missing outcome data. The assumption that data is at least missing at random mìght not hold. Multiple imputation methods will not remove or reduce the bias that occurs when missingness in the outcome depends on its true value, unless such missingness can be explained by measured variables.
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: "We withdrew five patients who, by therapist report and on independent evaluator assessment, developed worsening depression (two patients in the relaxation therapy group), manifested bipolar disorder (one patient in the IPTgroup), engaged in severe substance abuse (one patient in the IPT group), or violated protocol by obtaining outside treatment (one patient in the prolonged exposure group)."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	

			comment: although proportions of dropout are not significantly different between the intervention groups and participants with and without posttest data do not differ regarding their baseline characteristics, the possibility that data is MNAR cannot be ruled out completely. Reasons for dropout are not reported (beyond those quoted above) and might be related to the intervention. As an example, fear of systematic exposure in PE might have caused participants to dropout, especially if strong feelings of guilt or shame were associated with the traumatic experience. In view of the fact that there were only few sessions and little time to build a trustful relationship between therapist and participant, this possibility must be considered. Another example for reasons related to the intervention would be the worsening of symptoms which might not have been detected by the therapist or which might not have been considered severe enough to withdraw the participant yet motivated the participant to cease treatment. All in all, although an effort was made to reduce the risk of bias there are still some concerns
	Risk of bias judgement	**Some concerns**	although the methods that were used by the authours to handle missing data consider several potential sources of bias, they do not offer definite evidence that the results was not biased by missing outcome data. The assumption that data is at least missing at random is rather strong in this context and might not hold.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: " The widely used 30-item CAPS was used to assess PTSD symptoms on frequency and intensity scales. [...] Interrater reliability for frequency and severity is excellent for the intrusion, hyperarousal, and avoidance subscales (r values,.0.92). Each subscale has good internal consistency (Cronbach's alpha=0.87)"

	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "The independent evaluators achieved excellent interrater reliability on the CAPS (primary outcome measure; Shrout-Fleiss interclass reliability coefficient =0.93)"
			comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	comment: although the assessors were blind to the participants' condition, there was no blind assessment, since the participants (answering interview questions) were not blinded
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention. Thus, the probability that participants have answered according to their beliefs/expectations about the intervention effect decreases (as opposed to a passive control condition)
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN	
	Risk of bias judgement	**Some concerns**	although the interviewers were blinded, assessment was not completely blind as participants were probably aware of their treatment allocation which increases the risk that participants answer interview questions according to their expectations regarding the treatment efficacy or according to their beliefs about desired results (to please the investigators). The risk of bias due to lack of participant blinding, however, is lowered by the fact that the comparator was also an active intervention (and, in addition, sharing common elements of psychotherapeutic treatment). Some concerns, however, remain.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: "Clinicaltrials.gov identifier: NCT00739765."
			comment: examination of the history of changes. A comparison of study record versions (clinicaltrials.gov) indicates that authors adhered to their pre-specified intentions in all aspects that are presented and that are relevant for the result of interest

5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: see above. results were reported for all outcome measures of interest; for all time points of interest
5.3 ... multiple eligible analyses of the data?	PN	comment: see above. Authours report detailed justification for methods used to estimate means and SDs (ITT analysis using longitudinal mixed-effects models; multiple imputation), giving no rise to concerns
Risk of bias judgement	**Low**	The trial was prospectively registered at ClinicalTrials.gov. Examination of the history of changes and a comparison of this data with the reported information indicates that, all in all, researchers adhered to their pre-specified intentions.

Overall bias	**Risk of bias judgement**	**Some concerns**	no information on allocation concealment; ITT analysis with multiple imputation used; methods used to handle missing data do not offer definite evidence that the results was not biased by missing outcome data. risk of bias due to lack of participant blinding is lowered by the fact that the comparator was also an active intervention; some concerns, however, remain. The trial was pre-registered at ClinicalTrials.gov. Examination of the history of changes and a comparison of this data with the reported information indicates that, all in all, researchers adhered to their pre-specified intentions.

Unique ID	80	**Study ID**	870102	**Assessor**	R
Ref or Label	Markowitz 2015	**Aim**	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	PE	**Comparator**	REL	**Source**	Journal article(s) with results of the trial; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
Outcome	clinician-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomizatio n process	1.1 Was the allocation sequence random?	Y	Quote: "Eligible individuals who provided written informed consent for the treatment study were randomly assigned to receive prolonged exposure, IPT, or relaxation therapy, in a 4:4:3 ratio. Randomization followed a computer-generated program designed by the study's statistician, who had no patient contact. Randomization was stratified by presence of major depressive disorder (diagnosed according to the SCID, along with a score $20 on the 24-item Hamilton Depression Rating Scale [HAM-D] [33]) and implemented in blocks of random sizes (11 or 22)."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	comment: see table 2; table 3
	Risk of bias judgement	**Some concerns**	computer-generated random sequence; no information on allocation concealment; inspection of baseline characteristics gives no rise to concerns
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	Quote: "[...] and patients were reminded not to identify their therapy or therapist during evaluation."
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as it is hardly possible to achieve blinding when psychological interventions are implemented
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "Analyses followed the intention-to-treat principle. [...] Efficacy of the three treatments with respect to symptom severity

			was estimated based on longitudinal mixed-effects models (42) using multiple imputation for the missing values"
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	deviations (e.g. non-adherence by participants who did not attend all sessions) did not arise because of the trial context; ITT analysis with multiple imputation used
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "Attrition was 15% in the IPT group, 29% in the prolonged exposure group, and 34% in the relaxation therapy group (n.s.). Two patients from each treatment condition withdrew after randomization but before beginning therapy."

	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	Quote: " We compared participants with missing postrandomization data to those without missing data with respect to baseline characteristics. No comparisons between subjects with and without postrandomization assessment overall, or within treatment groups, were significantly different, and no differences approached clinically meaningful magnitude."
			"Foreach variable(scoreontheCAPS,thePosttraumaticStressScale–Self Report,theHAM-D,theSocialAdjustmentScale–Self-Report, thequality oflifemeasure,andtheInventory ofInterpersonal Problems), weused theMarkov chainMonteCarlo technique to obtain a monotone missing data pattern. We then applied a predictive mean-matching regression method separately for thethreetreatmentgroups(44);toincreasethelikelihoodthat the missing-at-random assumption is valid, in addition to the previousvaluesofthevariablebeingimputed,weusedallother symptom variables and baseline major depression status as predictors in predictive mean-matching regression. Fifty imputed data sets were generated."
			" The omnibus test assessing whether dropout depended on the interaction between depression status and treatment showed a p value of 0.15. Half of patients who had comorbid depression and were assigned to receive prolonged exposure dropped out: the odds ratio of prolonged exposure attrition with (50%) and without (5.6%) major depression was 17:1 (Table 4). Dropout among depressed patients in the prolonged exposure group tended to be higher than among depressed patients in the IPT group (p=0.086) and higher than dropout among nondepressed patients in the prolonged exposure group (p=0.006)."
			comment: although the methods that were used by the authours

			to handle missing data consider several potential sources of bias, they do not offer definite evidence that the results was not biased by missing outcome data. The assumption that data is at least missing at random mígnt not hold. Multiple imputation methods will not remove or reduce the bias that occurs when missingness in the outcome depends on its true value, unless such missingness can be explained by measured variables.
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: "We withdrew five patients who, by therapist report and on independent evaluator assessment, developed worsening depression (two patients in the relaxation therapy group), manifested bipolar disorder (one patient in the IPTgroup), engaged in severe substance abuse (one patient in the IPT group), or violated protocol by obtaining outside treatment (one
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	

			patient in the prolonged exposure group)."
			comment: although proportions of dropout are not significantly different between the intervention groups and participants with and without posttest data do not differ regarding their baseline characteristics, the possibility that data is MNAR cannot be ruled out completely. Reasons for dropout are not reported (beyond those quoted above) and might be related to the intervention. As an example, fear of systematic exposure in PE might have caused participants to dropout, especially if strong feelings of guilt or shame were associated with the traumatic experience. In view of the fact that there were only few sessions and little time to build a trustful relationship between therapist and participant, this possibility must be considered. Another example for reasons related to the intervention would be the worsening of symptoms which might not have been detected by the therapist or which might not have been considered severe enough to withdraw the participant yet motivated the participant to cease treatment. All in all, although an effort was made to reduce the risk of bias there are still some concerns
	Risk of bias judgement	**Some concerns**	although the methods that were used by the authours to handle missing data consider several potential sources of bias, they do not offer definite evidence that the results was not biased by missing outcome data. The assumption that data is at least missing at random is rather strong in this context and might not hold.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: " The widely used 30-item CAPS was used to assess PTSD symptoms on frequency and intensity scales. [...] Interrater reliability for frequency and severity is excellent for the intrusion, hyperarousal, and avoidance subscales (r values,.0.92). Each subscale has good internal consistency (Cronbach's alpha=0.87)"

	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "The independent evaluators achieved excellent interrater reliability on the CAPS (primary outcome measure; Shrout-Fleiss interclass reliability coefficient =0.93)" comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	comment: although the assessors were blind to the participants' condition, there was no blind assessment, since the participants (answering interview questions) were not blinded
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention. Thus, the probability that participants have answered according to their beliefs/expectations about the intervention effect decreases (as opposed to a passive control condition)
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN	
	Risk of bias judgement	**Some concerns**	although the interviewers were blinded, assessment was not completely blind as participants were probably aware of their treatment allocation which increases the risk that participants answer interview questions according to their expectations regarding the treatment efficacy or according to their beliefs about desired results (to please the investigators). The risk of bias due to lack of participant blinding, however, is lowered by the fact that the comparator was also an active intervention. Some concerns, however, remain.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: "Clinicaltrials.gov identifier: NCT00739765." comment: examination of the history of changes. A comparison of study record versions (clinicaltrials.gov) indicates that authors adhered to their pre-specified intentions in all aspects that are presented and that are relevant for the result of interest

	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: see above. results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: see above. Authours report detailed justification for methods used to estimate means and SDs (ITT analysis using longitudinal mixed-effects models; multiple imputation), giving no rise to concerns
	Risk of bias judgement	**Low**	The trial was prospectively registered at ClinicalTrials.gov. Examination of the history of changes and a comparison of this data with the reported information indicates that, all in all, researchers adhered to their pre-specified intentions.
Overall bias	**Risk of bias judgement**	**Some concerns**	no information on allocation concealment; ITT analysis with multiple imputation used; methods used to handle missing data do not offer definite evidence that the results was not biased by missing outcome data. risk of bias due to lack of participant blinding is lowered by the fact that the comparator was also an active intervention; some concerns, however, remain. The trial was pre-registered at ClinicalTrials.gov. Examination of the history of changes and a comparison of this data with the reported information indicates that, all in all, researchers adhered to their pre-specified intentions.

Unique ID	81	**Study ID**	870103	**Assessor**	R
Ref or Label	Markowitz 2015	**Aim**	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	IPT	**Comparator**	REL	**Source**	Journal article(s) with results of the trial; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
Outcome	clinician-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "Eligible individuals who provided written informed consent for the treatment study were randomly assigned to receive prolonged exposure, IPT, or relaxation therapy, in a 4:4:3 ratio. Randomization followed a computer-generated program designed by the study's statistician, who had no patient contact. Randomization was stratified by presence of major depressive disorder (diagnosed according to the SCID, along with a score $20 on the 24-item Hamilton Depression Rating Scale [HAM-D] [33]) and implemented in blocks of random sizes (11 or 22)."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	comment: see table 2; table 3
	Risk of bias judgement	**Some concerns**	computer-generated random sequence; no information on allocation concealment; inspection of baseline characteristics gives no rise to concerns
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	Quote: "[...] and patients were reminded not to identify their therapy or therapist during evaluation."
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as it is hardly possible to achieve blinding when psychological interventions are implemented
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "Analyses followed the intention-to-treat principle. [...] Efficacy of the three treatments with respect to symptom severity

			was estimated based on longitudinal mixed-effects models (42) using multiple imputation for the missing values"
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	deviations (e.g. non-adherence by participants who did not attend all sessions) did not arise because of the trial context; ITT analysis with multiple imputation used
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "Attrition was 15% in the IPT group, 29% in the prolonged exposure group, and 34% in the relaxation therapy group (n.s.). Two patients from each treatment condition withdrew after randomization but before beginning therapy."

	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	Quote: " We compared participants with missing postrandomization data to those without missing data with respect to baseline characteristics. No comparisons between subjects with and without postrandomization assessment overall, or within treatment groups, were significantly different, and no differences approached clinically meaningful magnitude."

"Foreach variable(scoreontheCAPS,thePosttraumaticStressScale–Self Report,theHAM-D,theSocialAdjustmentScale–Self-Report, thequality oflifemeasure,andtheInventory ofInterpersonal Problems), weused theMarkov chainMonteCarlo technique to obtain a monotone missing data pattern. We then applied a predictive mean-matching regression method separately for thethreetreatmentgroups(44);toincreasethelikelihoodthat the missing-at-random assumption is valid, in addition to the previousvaluesofthevariablebeingimputed,weusedallother symptom variables and baseline major depression status as predictors in predictive mean-matching regression. Fifty imputed data sets were generated."

" The omnibus test assessing whether dropout depended on the interaction between depression status and treatment showed a p value of 0.15. Half of patients who had comorbid depression and were assigned to receive prolonged exposure dropped out: the odds ratio of prolonged exposure attrition with (50%) and without (5.6%) major depression was 17:1 (Table 4). Dropout among depressed patients in the prolonged exposure group tended to be higher than among depressed patients in the IPT group (p=0.086) and higher than dropout among nondepressed patients in the prolonged exposure group (p=0.006)."

comment: although the methods that were used by the authours |

			to handle missing data consider several potential sources of bias, they do not offer definite evidence that the results was not biased by missing outcome data. The assumption that data is at least missing at random mígnt not hold. Multiple imputation methods will not remove or reduce the bias that occurs when missingness in the outcome depends on its true value, unless such missingness can be explained by measured variables.
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: "We withdrew five patients who, by therapist report and on independent evaluator assessment, developed worsening depression (two patients in the relaxation therapy group), manifested bipolar disorder (one patient in the IPTgroup), engaged in severe substance abuse (one patient in the IPT group), or violated protocol by obtaining outside treatment (one
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	

			patient in the prolonged exposure group)."
			comment: although proportions of dropout are not significantly different between the intervention groups and participants with and without posttest data do not differ regarding their baseline characteristics, the possibility that data is MNAR cannot be ruled out completely. Reasons for dropout are not reported (beyond those quoted above) and might be related to the intervention. As an example, fear of systematic exposure in PE might have caused participants to dropout, especially if strong feelings of guilt or shame were associated with the traumatic experience. In view of the fact that there were only few sessions and little time to build a trustful relationship between therapist and participant, this possibility must be considered. Another example for reasons related to the intervention would be the worsening of symptoms which might not have been detected by the therapist or which might not have been considered severe enough to withdraw the participant yet motivated the participant to cease treatment. All in all, although an effort was made to reduce the risk of bias there are still some concerns
	Risk of bias judgement	**Some concerns**	although the methods that were used by the authours to handle missing data consider several potential sources of bias, they do not offer definite evidence that the results was not biased by missing outcome data. The assumption that data is at least missing at random is rather strong in this context and might not hold.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: " The widely used 30-item CAPS was used to assess PTSD symptoms on frequency and intensity scales. [...] Interrater reliability for frequency and severity is excellent for the intrusion, hyperarousal, and avoidance subscales (r values,.0.92). Each subscale has good internal consistency (Cronbach's alpha=0.87)"

			Quote: "The independent evaluators achieved excellent interrater reliability on the CAPS (primary outcome measure; Shrout-Fleiss interclass reliability coefficient =0.93)"
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	
			comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	comment: although the assessors were blind to the participants' condition, there was no blind assessment, since the participants (answering interview questions) were not blinded
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention. Thus, the probability that participants have answered according to their beliefs/expectations about the intervention effect decreases (as opposed to a passive control condition)
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN	
	Risk of bias judgement	**Some concerns**	although the interviewers were blinded, assessment was not completely blind as participants were probably aware of their treatment allocation which increases the risk that participants answer interview questions according to their expectations regarding the treatment efficacy or according to their beliefs about desired results (to please the investigators). The risk of bias due to lack of participant blinding, however, is lowered by the fact that the comparator was also an active intervention. Some concerns, however, remain.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: "Clinicaltrials.gov identifier: NCT00739765."
			comment: examination of the history of changes. A comparison of study record versions (clinicaltrials.gov) indicates that authors adhered to their pre-specified intentions in all aspects that are presented and that are relevant for the result of interest

	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: see above. results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: see above. Authours report detailed justification for methods used to estimate means and SDs (ITT analysis using longitudinal mixed-effects models; multiple imputation), giving no rise to concerns
	Risk of bias judgement	**Low**	The trial was prospectively registered at ClinicalTrials.gov. Examination of the history of changes and a comparison of this data with the reported information indicates that, all in all, researchers adhered to their pre-specified intentions.
Overall bias	**Risk of bias judgement**	**Some concerns**	no information on allocation concealment; ITT analysis with multiple imputation used; methods used to handle missing data do not offer definite evidence that the results was not biased by missing outcome data. risk of bias due to lack of participant blinding is lowered by the fact that the comparator was also an active intervention; some concerns, however, remain. The trial was pre-registered at ClinicalTrials.gov. Examination of the history of changes and a comparison of this data with the reported information indicates that, all in all, researchers adhered to their pre-specified intentions.

Unique ID	86	Study ID	920101	Assessor	R
Ref or Label	McDonagh 2005	**Aim**	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	PCT	**Comparator**	WL	**Source**	Journal article(s) with results of the trial

Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomizatio n process	1.1 Was the allocation sequence random?	NI	Quote: "Following the initial laboratory evaluation, women were randomly assigned to one of the following three conditions for 14 weeks: CBT, PCT, or WL. When it became clear that the dropout rate was greater for CBT, we changed the random assignment process to increase the chance of assignment to CBT."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "There were no differences among the three groups on any study measures or demographic characteristics (ps > .05)."
	Risk of bias judgement	**Some concerns**	no information on random sequence generation and allocation concealment; no differences between groups on any measures or demographic variables at baseline.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	Quote: " Participants assigned to the WL were told that they could receive their choice of the two treatments in about 14 weeks, after completing the post-WL assessment."
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as it is hardly possible to achieve blinding when psychological interventions are implemented
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	Quote: " For both CBT and PCT, adherence was excellent at 87.80% (SD = 7.19) and 88.50% (SD = 6.96), respectively. Interrater reliability was good (k= .80)." comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias. All reported deviations could occur outside the trial context

	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "We first conducted intention to treat analyses using data on all who were randomized to treatment.These analyses were of necessity very conservative ones, because we lacked postdropout assessments. Therefore, admission data were carried forward to subsequent assessment points for participants who dropped out of treatment."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	therapists and participants were probably aware of the assigned condition; no deviations that arose because of the experimental context; ITT analysis.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "The dropout rate for the study was 23%, with a rate of 41% (12 of 29) for CBT, 9% (2 of 22) for PCT, and 13% (3 of 23) for WL" comment: 9% missing posttest data in the PCT group; 13% in the WL group
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis reported
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: "Analysis of the trauma measures that compared participants who completed the study with those who dropped out revealed differences in the frequency of childhood physical abuse (Fisher's exact test, $p < .02$), perceived threat during the worst CSA event, $x^2(1, N = 74) = 4.63$, $p < .05$, and physical injury during the worst CSA event, $x^2(1, N = 74) = 5.94$, $p < .02$, with the dropout group having higher scores on all of these variables than the completer group."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	

			"Because of the higher dropout rate in CBT than in PCT or WL in this study, comparisons between CBT and the other two groups are less scientifically sound than those comparing PCT and WL, in which dropouts were few."
			comment: equal proportions of dropout between both groups. the overall amount of missing data is relatively small. However, reasons for dropout (n=5) are not reported. The information quoted above suggests that dropouts and completers differed with respect to their health status which would lead to bias. It should be noted, however, that this analysis included the CBT group, the group with the largest number of dropouts (n= 12; = 41%). It is unknown whether PCT/WL dropouts differed from PCT/WL completers on any variables. All in all, there are some concerns but there is no evidence for a biased result
	Risk of bias judgement	**Some concerns**	equal proportions of dropout between both groups; the overall amount of missing data is relatively small; however, reasons for dropout are not reported.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The CAPS (Weathers et al., 2001) provides a standardized method for making current and lifetime DSM–IV diagnosis of PTSD. The scale also measures the intensity and frequency of the 17 individual PTSD symptoms." comment: The administered scale (CAPS) is a validated PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points

	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "A separate group of female clinicians, who were blind to treatment condition and who had no other role in the study conducted the four CAPS interviews." comment: although the assessors were blind to the participants' condition, there was no blind assessment, since the participants (answering interview questions) were probably aware of their treatment allocation
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: although the interviewers were blinded, there is still a risk of bias. Because the comparator is a passive control condition the risk is higher that participants might have answered questions according to their beliefs/expectations regarding their assigned condition. This risk can not be fully eliminated by assessment by a blinded clinician.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	although the interviewers were blinded, assessment was not completely blind: participants were probably aware of their treatment allocation which increases the risk that participants answered interview questions according to their expectations regarding the treatment efficacy or according to their beliefs about desired results (e.g. to please the investigators). This risk is particularly high as the comparator was a no-treatment condition.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: results from both ITT and PP analysis are reported; generally, the risk of bias due to multiple eligible analyses of the data is low as the results assessed here are raw values (means, SDs)

	Risk of bias judgement	Some concerns	no information regarding pre-specified analysis plan; results were reported for all outcome measures of interest; for all time points of interest; ITT results as well as results for completers are reported; generally, the risk of bias due selection of results based on multiple eligible analyses is low as the results assessed here are raw values (means, SDs). All in all, difficult to assess due to lack of information; hence, some concerns cannot be eliminated
Overall bias	Risk of bias judgement	High	no information on random sequence generation and allocation concealment; no differences between groups on any measures or demographic variables at baseline. equal proportions of dropout between both groups; the overall amount of missing data is relatively small; however, reasons for dropout are not reported. risk of bias due to participants' knowledge of intervention. no sufficient information regarding pre-specified analysis plan.

Unique ID	89	Study ID	960101	Assessor	R
Ref or Label	Mitchell 2014	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	MBI	Comparator	WL	Source	Journal article(s) with results of the trial
Outcome	self-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the	1.1 Was the allocation sequence random?	Y	Quote: "The principal investigator (PI) used the Microsoft Excel random numbers function to assign participants to groups."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	

randomizatio n process	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "There were no significant differences in ethnicity, education, whether they previously had practiced yoga, or PCL, STAI, or CES-D baseline scores (all p > .05)."
	Risk of bias judgement	**Some concerns**	computer-generated random number sequence; no information regarding allocation concealment; no baseline differences between groups.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	Y	Quote: "Neither the group instructors nor participants were blinded to the randomization, as this would not have been feasible."
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	PY	Quote: "The full intent-to-treat (ITT) sample (n=38) was included in the analyses. All available data were used in the growth curve models (Muth´en & Muth´en, 1998-2010)."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding of participants or instructors; ITT analysis.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote from Figure 1: "Allocated to Yoga (20), Lost to Follow-Up (3), Withdrew (3)" "Allocated to control (18), Lost to Follow-Up (2), Withdrew (4)" comment: amount of missing posttest data 30% (n=6) in the MBI group; 33.3% in the WL group

3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis reported
3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: "There were no significant differences between study completers and noncompleters in terms of age, t(36) =−1.85, p=.073), BMI, t(36)=0.34, p= .738, education χ2 (3, N=38) = 5.82, p = .152, ethnicity χ2 (3, N = 38) = 1.14, p = .768, or the proportion who had taken a yoga class before χ2 (1, N= 38)=0.05, p= .852. There was a marginally significant difference in baseline PCL scores, with noncompleters having higher scores (M = 59.20) than did completers (M = 49.83); t(31) =2.08, p= .046."
3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	"As noted above, noncompleters had marginally significantly higher PCL scores than did study completers. An equal number of participants, however, dropped out of the yoga and control groups. No adverse reactions were noted in either group" "Our participants did not report worsened symptoms or other adverse reactions as a result of the intervention, and it is important to have demonstrated this empirically." comment: equal proportions of dropout between both groups; the overall amount of missing data is relatively large (one third); no adverse events were documented, however, noncompleters showed slightly more severe symptoms at baseline; reasons for dropout are not reported. All in all, although proportions of dropout were equal it is possible that missingness in the outcome depended on its true value
Risk of bias judgement	**High**	equal proportions of dropout between both groups; the overall amount of missing data is relatively large (one third); no adverse events were documented; noncompleters showed slightly more severe PTSD symptoms at baseline; reasons for dropout not reported.

Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The PTSD Checklist-Civilian (PCL-C)isa17-item measure of DSM-IV (APA, 2000). [...] In the current study, Cronbach's α at baseline was .87." comment: The administered scale (PCL) is a validated PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "In the assessment control condition, participants met once per week for 12 weeks in groups of 4–5 to complete the same weekly questionnaires as yoga participants." comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "Neither the group instructors nor participants were blinded to the randomization, as this would not have been feasible." comment: the PCL is a self-report questionnaire; so the 'assessors' were the participants themselves - who were aware of their group allocation status
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: assessment could have been influenced because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is high, as the comparator is a passive control condition (here: WL).

	Risk of bias judgement	**High**	Risk of bias due to knowledge of the intervention as participants were not blind to their condition. It is particularly high as the comparator was a no-treatment condition.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	Quote: "The PC-PTSD is a 4-item measure that was used to assess past-month PTSD symptoms during the telephone screening" "The PTSD Symptom Scale-Interview (PSS-I) was used to diagnose PTSD prior to randomization" comment: the PSS-I as well as the PC-PTSD were also administered (at least at baseline). It is not clearly stated that there were no assessments after treatment using those measures; and no justification is offered why those other PTSD measures were only administered at baseline. This might raise some concerns.
	5.3 ... multiple eligible analyses of the data?	PN	comment: the risk of bias due selection of results based on multiple eligible analyses is lowered by the fact that results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for all outcome measures of interest; for all time points of interest; results for the ITT sample reported; slight concerns regarding selection of results based on results from multiple eligible outcome measures; generally, the risk of bias due selection of results based on multiple eligible analyses is low as the results assessed here are raw values (means, SDs). All in all, difficult to assess due to lack of information; hence, some concerns cannot be eliminated

Overall bias	Risk of bias judgement	High	computer-generated random number sequence; no information regarding allocation concealment; no baseline differences between groups. no blinding of participants or instructors; ITT analysis. , although proportions of dropout were equal it is possible that missingness in the outcome depended on its true value (no reasons for dropout reported; dropouts higher baseline PCL scores). Risk of bias due to knowledge of the intervention as participants were not blind to their condition; it is particularly high as the comparator was a no-treatment condition. not enough information regarding pre-specified analysis plan to reliably assess the risk of bias due to selection of the reported result.

Unique ID	90	Study ID	970101	Assessor	R
Ref or Label	Monson 2006	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	CPT	Comparator	WL	Source	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	PY	Quote: "Eligible participants were randomized to receive the treatment immediately or to wait for 10 weeks to receive the treatment (10 weeks was equivalent to the ideal 6 weeks of twice weekly sessions and the 1-month follow-up period for those in the CPT condition). The study biostatistician provided the participants' condition assignment to the study coordinator."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "Table 1 provides descriptive information about the sample overall and by condition. There were no statistically significant

			differences between the two conditions in baseline characteristics."
	Risk of bias judgement	**Some concerns**	probably allocation sequence was random; no information on allocation concealment; no baseline differences.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as it is hardly possible to achieve blinding when psychological interventions are implemented
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "Primary analyses were performed according to the intention-to-treat principle; data from all participants were used regardless of their treatment completion."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	therapists and participants were probably necessarily aware of group status; ITT analysis used.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "The overall dropout rate was 16.6% (20% from CPT, 13% from the wait-list condition)." comment: see figure 1: out of n=30 CPT participants n=24 were post-assessed; out of n=30 WL participants n=27 were post-assessed

	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	Quote: "Primary analyses were performed according to the intention-to-treat principle; data from all participants were used regardless of their treatment completion. We also examined data from participants who completed the treatment (50 of 60 participants), and the results were highly consistent with the results found in the intention-to-treat sample" comment: no analysis correcting for bias or sensitivity analysis reported. The approach described in the quote offers no reliable evidence that the result was not biased by missing data
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NI	Quote: " To better understand these findings, we inspected the individual outcomes of the participants who dropped out. Two of the six participants who dropped out of CPT improved with less than the full course of therapy. These findings, in tandem with other recent reports (Resick, Williams, Orazem, & Gutner, 2005), reinforce that treatment dropout is not necessarily an indicator of poor tolerance of therapy."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	comment: the overall dropout rate was moderate (16.6%). the information quoted (3.1; 3.3) offers no evidence that the result was not biased. proportions of dropouts were approximately equal in both groups, but posttest data is missing for n=9 participants - two of those dropouts showed symptom improvement before cessassion of treatment but there is no information regarding the remaining seven dropouts. Reasons for dropout are not reported. It was not tested (or not reported) whether completers differed from dropouts on demographic or clinical variables at baseline. There is no evidence for bias, but it is unclear whether missingness in the outcome depended on its value. There are concerns.

	Risk of bias judgement	**High**	results from ITT analysis and complerers analysis are consistent. the overall dropout rate was moderate (16.6%). proportions of dropouts did not differ between groups. posttest data missing for n=9; two CPT dropouts showed symptom improvement before cessassion of treatment, but no information regarding the remaining 7 dropouts. Reasons for dropout not reported; not reported whether completers differed from dropouts on any demographic or clinical variables at baseline. There is no evidence for bias, however, very limited information is provided which raises concerns.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: " The Clinician-Administered Posttraumatic Stress Disorder Scale (CAPS; Blake et al., 1995) was used to determine PTSD diagnostic status and severity. The CAPS is a widely used and validated clinician interview for the assessment of PTSD (Weathers, Keane, & Davidson, 2001)." " All SCID-P and CAPS assessments were audiotaped; 10% of the SCID-P and 7.5% of the CAPS administered were evaluated by an independent doctoral-level clinical psychologist for reliability. The intraclass correlation for PTSD severity on the CAPS showed excellent agreement (rs = .72 to .99 across symptom clusters)."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "The independent clinician assessors were blinded to condition assignment and participants were instructed to not disclose their condition assignment to them." comment: although the assessors were blind to the participants' condition, there was no blind assessment, since the participants (answering interview questions) were probably aware of their treatment allocation

	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: although the interviewers were blinded, there is still a risk of bias. Because the comparator is a passive control condition the risk is higher that participants might have answered questions according to their beliefs/expectations regarding their assigned condition. This risk can not be fully eliminated by assessment by a blinded clinician.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	despite of the blinding of interviewers there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. The risk is particularly relevant as the comparator was a no-treatment waitlist condition.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest (self- and clinician-rated PTSD measures); for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: both ITT and completers data reported; generally, the risk of bias due to multiple eligible analyses of the data is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure(s) of interest; for all time points of interest; results for the ITT sample reported; generally, the risk of bias due selection of results based on multiple eligible analyses is lowered by the fact that the results assessed here are raw values (means, SDs). All in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are some concerns.

Overall bias	Risk of bias judgement			High	probably the allocation sequence was random; no information on allocation concealment; no baseline differences. ITT analysis used. There is no evidence for bias due to missingness in the outcome, however, very limited information is provided which raises concerns. despite of the blinding of interviewers there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition. not enough information regarding pre-specified analysis plan to reliably assess the risk of bias due to selection of the reported result.

Unique ID	92	Study ID	1000101	Assessor	R
Ref or Label	Morath 2014	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	NET	Comparator	WL	Source	Journal article(s) with results of the trial; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: " An independent person randomly assigned individuals with PTSD to either a treatment condition (NET) or the WLC group using permuted blocks of variable length. The capacity of the therapists who carried out NET was the criterion for the lengths of the blocks – that is, if the next therapist to be sent patients had k free therapy slots, then k of the next $2k$ participants would be randomly assigned to the NET group and referred to this therapist, while the other k participants would be assigned to the WLC group, using a shuffled set of envelopes which were opened only after a new participant was included in the study. Diagnosticians
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	PY	

			were not aware of which participants were allocated to which group."
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "Groups presented with very similar socio-demographic and clinical characteristics prior to treatment." " Pre-therapy PTSD symptom severity (CAPS score) did not differ between the NET and the WLC groups (table 2)." " Before therapy, groups did not differ in basal DNA strand breaks (table 2)."
	Risk of bias judgement	**Low**	randomization using permuted blocks; reported information indicates that there was allocation concealment; no baseline differences on relevant variables between groups.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as it is hardly possible to achieve blinding when psychological interventions are implemented
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: " Linear mixed models were calculated to analyse the primary and secondary outcomes of changes in DNA breakage and repair as well as changes in PTSD symptom severity (CAPS score)."comment: mITT analysis included all participants randomized
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	therapists and participants were probably necessarily aware of the assigned condition; mITT analysis included all participants randomized

Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	comment: see online supplement (participant flow): n=4 out of n=19 in the NET group were lost to post-assessment; n=5 out of n=19 in the WL group were lost to post-assessment [missing posttest data: 23.7% overall; 21% in NET; 26.3% in WL]
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis reported
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote from participant flow (online supplement): NET: "Lost to Post-test (n=4) (stopped treatment because of alcohol problems n=2, moved to unknown address n=1, in prison n=1)"
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	WL: "Lost to Post-test (n=5) (moved to unknown address)" comment: n= 2 NET cases [5% of the total sample; 10.5% of NET group] where missingness in the outcome might have depended on its true value. Proportions of dropout are equal between groups; no information regarding differences between participants with and without posttest data on sociodemographic or clinical variables at baseline. Given that reasons for dropout were unrelated in most cases the risk of bias is not rated high. There are, however, some concerns
	Risk of bias judgement	**Some concerns**	missing posttest data: 23.7% overall, equal proportions; no analysis correcting for bias or sensitivity analysis; Given that reasons for dropout were unrelated in most cases the risk of bias is not rated high. There are, however, some concerns.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: " Measures of outcome were changes in DNA breakage and repair 4 months and 1 year after the end of treatment with NET (primary endpoint) and the diagnosis of PTSD and the change of its severity score according to CAPS (secondary endpoint)."

			comment: The CAPS is a gold-standard, validated PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "Diagnosticians were not aware of which participants were allocated to which group. Blinded diagnosticians conducted post-test and follow-up interviews."
			comment: although the assessors were blind to the participants' condition, there was no blind assessment, since the participants (answering interview questions) were probably aware of their treatment allocation
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: although the interviewers were blinded, there is still a risk of bias. Because the comparator is a passive control condition the risk is higher that participants might have answered questions according to their beliefs/expectations regarding their assigned condition. This risk can not be fully eliminated by assessment by a blinded clinician.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	despite of the blinding of interviewers there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. The risk is particularly relevant as the comparator was a no-treatment waitlist condition.

Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: " The trial was registered at clinicaltrials.org, NCT01206790." comment: The trial was pre-registered (clinicaltrials.gov). examination of the history of changes and the comparison of study record versions indicates that authors adhered to their pre-specified intentions in all aspects that are presented and that are relevant for the result of interest (one change: one exclusion criterion was changed for undocumented reasons [participants with current alcohol/drug abuse no longer excluded; instead participants with chronic inflammatory deseases])
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: results are reported for all participants randomized (ITT sample); generally, the risk of bias due selection of results based on multiple eligible analyses is lowered by the fact that results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Low**	The trial was prospectively registered at ClinicalTrials.gov. A comparison of this data with the reported information indicates that, all in all, researchers adhered to their pre-specified intentions. The risk of bias is therefore low.
Overall bias	**Risk of bias judgement**	**High**	randomization using permuted blocks; reported information indicates that there was allocation concealment; no baseline differences on relevant variables between groups. mITT analysis included all participants randomized. Given that the dropout rate was equal in both groups and that reasons for dropout were unrelated in most cases the risk of bias is not rated high -there are, however, some concerns. The trial was pre-registered at ClinicalTrials.gov; examination of information gives no rise to concerns.

Unique ID	93	Study ID	1040101	Assessor	R
Ref or Label	Neuner 2004	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	NET	Comparator	Psychoeducation	Source	Journal article(s) with results of the trial
Outcome	self-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	PY	Quote: "Each participant was randomly assigned (using a dice) to one of three treatment groups: narrative exposure therapy, supportive counseling, or psychoeducation only."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	" The randomization procedure resulted in different group sizes"
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	Quote: "The randomization procedure resulted in different group sizes. There were no systematic group differences in any of the sociodemographic variables, as confirmed by Fisher's exact test and analysis of variance (ANOVA)." "As the randomization procedure resulted in different baseline levels for some measures, repeated measures ANOVAs [...]" "The randomization procedure resulted in different group sizes." "As the randomization procedure resulted in different baseline levels for some measures [...]" comment: no substantial excess in statistically sign. differences; the small number of significant differences should be compatible with chance; however, differences in intervention group sizes,

			compared with the intended allocation ratio, which might raise some concerns; the difference in pretreatment PTSD scores might lead to a biased effect size estimate (SMD of posttreatment scores)
	Risk of bias judgement	**Some concerns**	randomization by throwing a dice; no information regarding allocation concealment; small number of significant differences which should be compatible with chance, however, randomization resulted in differences in group sizes, compared to the intended allocation ratio.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	Quote: "The respondents were instructed not to inform the interviewers or the trained researchers about the type of treatment or the number of sessions they had received."
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	comment: therapists and participants (see quote) were probably necessarily aware of the assigned condition as it is hardly possible to achieve blinding when psychological interventions are implemented
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	Quote: "No major deviations from treatment protocol were detected."

			comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: " All participants who were randomized to treatment, including the 1 who refused participation in the narrative exposure therapy group and the 2 dropouts in the supportive counseling group, were included in the analyses." "To maximize use of information in this study with a small sample size, missing data were estimated with a restricted maximum likelihood procedure."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	therapists and participants were most likely aware of group status; mITT analysis using a restricted maximum likelihood procedure to estimate missing data.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	Quote: " In the narrative exposure therapy group 1 patient refused to participate; all other patients agreed. Only in the supportive counseling group did any patients fail to complete the full treatment; 2 patients in this group discontinued treatment." "The fact that there was no dropout in the narrative exposure therapy group is noteworthy." comment: see Table 2. In table 2 it is reported that posttest data was available for n= 15 out of 17 NET participants. No further

			information reported regarding this missing data and no reason for refusal of 1 NET participant reported. No missing data in the Psychoeducation group.
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	missing posttest data for n=2 NET participants; no missing data in the PsEd group.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	NI	Quote: " The frequency and severity of PTSD symptoms was assessed with the PDS (Foa, 1995), modified in the translation procedure to simplify the frequency rating of symptoms." comment: the PDS is a validated PTSD measure likely to be sensitive to treatment effects; however, no details are reported on modifications of the instrument in this study
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "The respondents were instructed not to inform the interviewers or the trained researchers about the type of treatment or the number of sessions they had received."

			comment: the PDS is a self-report questionnaire; hence the 'assessors' (the participants themselves) were probably aware of their intervention
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: assessment could have been influenced because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. This risk is lowered by the fact that the comparator was also an active intervention; however, there are substantial differences between the two interventions, e.g. regarding the number of sessions (NET= 4 sessions, PsEd=1) or trauma exposure. All in all (no blinding, self-report, differences between interventions), there is a risk of bias favouring NET
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	A modified version of the PDS was used (no details reported); no blinding of participants, outcome measured by self-report, differences between interventions that increase the risk of knowing one's group status; all in all, there is a risk of bias.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: results are reported for all participants randomized (ITT sample); generally, the risk of bias due selection of results based on multiple eligible analyses is lowered by the fact that results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure(s) of interest; for all time points of interest; results for the ITT sample reported; generally, the risk of bias due selection of results based on multiple eligible analyses is lowered by the fact that the results assessed here are raw values (means, SDs). All in all, it is difficult to assess the risk of

			bias in this domain due to lack of information; There are some concerns.
Overall bias	**Risk of bias judgement**	**High**	randomization by throwing a dice; no information regarding allocation concealment; small number of significant differences which should be compatible with chance, however, randomization resulted in differences in group sizes, compared to the intended allocation ratio. therapists and participants were most likely aware of group status; no deviations that arose because of the trial context; mITT analysis. missing posttest data for n=2 NET participants; no missing data in the PsEd group. A modified version of the PDS was used (no details reported); no blinding of participants, outcome measured by self-report, differences between interventions that increase the risk of knowing one's group status; all in all, there is a risk of bias. not enough information on pre-specified analysis plan to reliably assess the risk of bias due to selection of reported results. due to a difference in pretreatment PTSD severity between both groups (NET higher mean score) the SMD might be biased

Unique ID	94	**Study ID**	1070101	**Assessor**	R
Ref or Label	Nidich 2018	**Aim**	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	PE	**Comparator**	MBI	**Source**	Journal article(s) with results of the trial; Trial protocol; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)

Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "Stratified block randomisation, stratified on gender and years since military service release, was used to assign participants to their study treatment. Treatment groups were matched on number of treatment visits (12 sessions), length of each session (90 minutes), and duration of treatment (12 weeks)."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	"Allocation concealment was achieved by an off-site coinvestigator (JS), who randomly assigned each participant to a treatment group and informed the study coordinator (EM) of the treatment assignments (who then notified the participant)."
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	Quote: "Groups were compared on baseline and demographic variables, with ANOVA for continuous variables and χ^2 tests for categorical variables." comment: results of significance test not reported, but reported baseline characteristics (table 1) reveal no major differences between groups
	Risk of bias judgement	**Low**	stratified block randomization; allocation concealment; no substantial differences at baseline
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	Quote: "This single-blind, three-arm randomised controlled trial [...]"
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	"Additionally, all participants were asked not to divulge their treatment assignment." comment: blinding of participants and therapists is not feasible in the context of the intervention study

	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "All participants who were randomly assigned were included in the analyses, following the intention-to-treat principle. We did multiple imputations to include missing values at each time point, including the interim post-test visits, using the SAS software MI Procedure (Markov Chain Monte Carlo method; [SAS version 9.1.3])."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	blinding of participants and therapists is not feasible; ITT analyses conducted
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "166 (81%) participants completed the final 3-month posttest (53 (78%) of 68 for TM, 57 (84%) of 68 for PE, 56 (85%) of 66 for HE). All 202 eligible patients randomly assigned to treatment were included in the intention-totreat analyses regardless of treatment dropout or missing post-test data."

	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	Quote: " However, we also used the LOCF imputation method to estimate treatment effects in a sensitivity analysis because multiple imputation might yield less conservative estimates [...]" "A secondary analysis of change based on the LOCF method of imputing data yielded slightly more conservative results in terms of reductions in PCL-M and PHQ-9 scores compared with the intention-to-treat analysis based on multiple imputation; however, the LOCF method yielded similar results regarding statistical significance of between-group differences (appendix)." "Missing final posttest scores were imputed on the basis of change from baseline to non-missing interim and post-tests, taking into account correlations between non-missing values for primary and all secondary outcome scores at baseline, interim post-tests (PCL-M and PHQ-9 only) and final post-test." comment: Imputing the outcome variable through methods such as 'last-observation-carried-forward' or via multiple imputation based only on intervention group, should not be assumed to correct for bias due to missing outcome data. In this case MI was done taking additional variables/correlations into account, reducing the risk of bias. However, no sensitivity analyses was done showing that results are little changed under a range of plausible assumptions about the relationship between missingness in the outcome and its true value. The evidence is not considered sufficient to eliminate a risk of bias
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	Quote: "There were no treatment-related adverse events for any of the three treatments. The numbers of serious adverse events reported during the trial were not significantly different among treatment groups. There were three serious adverse events in the TM group (two suicide attempts, one death [non-suicidal]), two in the PE group (one drug overdose, one illness), and two in the HE
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	

			group (two psychiatric hospitalisations)."
			comment: reasons for declined/missed post-assessments of participants are unknown or not reported in detail for most participants with missing posttest data which increases the risk of bias. Reasons may be related to the treatment or the participants' health status. Proportions of missing posttest data did not differ between groups and there were no differences in adverse events. For details regarding the analysis see above. Multiple Imputation was done taking additional variables/correlations into account, reducing the risk of bias due to missing outcome data. However, their sensitivity analysis does not show that results are little changed under a range of plausible assumptions about the relationship between missingness in the outcome and its true value. The evidence indicates that findings are likely to be robust (MI/ LOCF) but evidence is not considered sufficient to eliminate all risk of bias due to missing data. There are some concerns
	Risk of bias judgement	Some concerns	equal proportions of missing outcome data between groups; reasons for refusal/non-attendance of post-assessments are unknown, so it is possible that missingness depends on its true value; multiple imputation was done on the basis of change from baseline to non-missing interim and post-tests, taking into account correlations between non-missing values for primary and all secondary outcome scores at baseline, interim post-tests and final post-test data. This approach is likely to lead to a lower risk of bias compared to MI based only on intervention group or compared to the LOCF method. Results were compared to results from LOCF procedure; similar results regarding significance of between-group differences. All in all, there are concerns.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The CAPS questionnaire was administered at baseline and at 3-months posttest. The research assistants administering the CAPS received training with an expert in CAPS administration

			(PH). The CAPS interview Cronbach's α ranges from 0.87 to 0.94, which indicates adequate internal consistency."
4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN		Quote: " Administration of the CAPS interview was supervised by two psychologists (TR and PH, study pyschologists)."
			comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
4.3 Were outcome assessors aware of the intervention received by study participants?	PY		Quote: "All baseline and post-test data were collected by two research assistants (MG and AR in the acknowledgments), masked to treatment assignment and uninvolved in any aspect of treatment delivery. Additionally, all participants were asked not to divulge their treatment assignment."
			comment: although the assessors were blind to the participants' condition, there was no blind assessment, since the participants (answering interview questions) were probably aware of their treatment allocation
4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY		comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention. Thus, it is less likely that participants might have answered according to their beliefs/expectations about the intervention effect.
4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN		
Risk of bias judgement	**Some concerns**		although the interviewers were blinded, assessment was not completely blind (participants might have been aware of their treatment allocation), increases the risk that participants answered interview questions according to their expectations regarding the treatment efficacy or according to their beliefs about desired results (to please the investigators). The risk of bias due to lack of

			participant blinding, however, is significantly lowered by the fact that the comparator was also an active treatment condition.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	Quote: "This study is registered with ClinicalTrials.gov, number NCT01865123." "Rutledge T, Nidich S, Schneider R, et al. Design and rationale of a comparative effectiveness trial evaluating transcendental meditation against established therapies for PTSD. Contemp Clin Trials 2014; 39: 50–56." comment: The trial was pre-registered (clinicaltrials.gov). examination of the history of changes and the comparison of the report with the published study protocol (published 2014, investigators were recruiting between 2013-2016) indicates that authors adhered to their pre-specified intentions in all aspects that are presented and that are relevant for the result of interest
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	comment: see above, all details reported regarding measurements are consistent with the trial protocol published before analysis
	5.3 ... multiple eligible analyses of the data?	N	comment: see above, all details reported regarding data analysis producing the result of interest are consistent with the trial protocol (published before analysis)
	Risk of bias judgement	**Low**	The trial was prospectively registered (clinicaltrials.gov) and a published study protocol is available. examination of the history of changes and comparison of the report with the published study protocol indicates that authors adhered to their pre-specified intentions in all aspects that are presented and that are relevant for the result of interest

Overall bias	Risk of bias judgement	Some concerns	equal proportions of missing outcome data between groups, reasons for refusal/non-attendance of post-assessments are unknown, ITT analyses using multiple imputation (taking several variables/correlations into account) and LOCF methods yielded similar results; some concerns remain. risk of bias due to lack of participant blinding is lowered by the fact that the comparator was also an active treatment condition, but there are concerns. The trial was pre-registered (clinicaltrials.gov) and a published study protocol is available; reported information is consistent with their pre-specified intentions, therefore risk of bias due to selective reporting is low.

Unique ID	95	Study ID	1070102	Assessor	R
Ref or Label	Nidich 2018	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	PE	Comparator	Psychoeducation	Source	Journal article(s) with results of the trial; Trial protocol; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "Stratified block randomisation, stratified on gender and years since military service release, was used to assign participants to their study treatment. Treatment groups were matched on number of treatment visits (12 sessions), length of each session (90 minutes), and duration of treatment (12 weeks)."

"Allocation concealment was achieved by an off-site coinvestigator (JS), who randomly assigned each participant to a |
| | 1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions? | Y | |

			treatment group and informed the study coordinator (EM) of the treatment assignments (who then notified the participant)."
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	Quote: "Groups were compared on baseline and demographic variables, with ANOVA for continuous variables and χ² tests for categorical variables." comment: results of significance test not reported, but reported baseline characteristics (table 1) reveal no major differences between groups
	Risk of bias judgement	**Low**	stratified block randomization; allocation concealment; no substantial differences at baseline
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	Quote: "This single-blind, three-arm randomised controlled trial [...]"
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	"Additionally, all participants were asked not to divulge their treatment assignment." comment: blinding of participants and therapists is not feasible in the context of the intervention study
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "All participants who were randomly assigned were included in the analyses, following the intention-to-treat principle. We did multiple imputations to include missing values at each time point, including the interim post-test visits, using the SAS software

			MI Procedure (Markov Chain Monte Carlo method; [SAS version 9.1.3])."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	blinding of participants and therapists is not feasible; ITT analyses conducted
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "166 (81%) participants completed the final 3-month posttest (53 (78%) of 68 for TM, 57 (84%) of 68 for PE, 56 (85%) of 66 for HE). All 202 eligible patients randomly assigned to treatment were included in the intention-totreat analyses regardless of treatment dropout or missing post-test data."

	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	Quote: " However, we also used the LOCF imputation method to estimate treatment effects in a sensitivity analysis because multiple imputation might yield less conservative estimates [...]" "A secondary analysis of change based on the LOCF method of imputing data yielded slightly more conservative results in terms of reductions in PCL-M and PHQ-9 scores compared with the intention-to-treat analysis based on multiple imputation; however, the LOCF method yielded similar results regarding statistical significance of between-group differences (appendix)." "Missing final posttest scores were imputed on the basis of change from baseline to non-missing interim and post-tests, taking into account correlations between non-missing values for primary and all secondary outcome scores at baseline, interim post-tests (PCL-M and PHQ-9 only) and final post-test." comment: Imputing the outcome variable through methods such as 'last-observation-carried-forward' or via multiple imputation based only on intervention group, should not be assumed to correct for bias due to missing outcome data. In this case MI was done taking additional variables/correlations into account, reducing the risk of bias. However, no sensitivity analyses was done showing that results are little changed under a range of plausible assumptions about the relationship between missingness in the outcome and its true value. The evidence is not considered sufficient to eliminate a risk of bias
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	Quote: "There were no treatment-related adverse events for any of the three treatments. The numbers of serious adverse events reported during the trial were not significantly different among treatment groups. There were three serious adverse events in the TM group (two suicide attempts, one death [non-suicidal]), two in the PE group (one drug overdose, one illness), and two in the HE
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	

			group (two psychiatric hospitalisations)."
			comment: reasons for declined/missed post-assessments of participants are unknown or not reported in detail for most participants with missing posttest data which increases the risk of bias. Reasons may be related to the treatment or the participants' health status. Proportions of missing posttest data did not differ between groups and there were no differences in adverse events. For details regarding the analysis see above. Multiple Imputation was done taking additional variables/correlations into account, reducing the risk of bias due to missing outcome data. However, their sensitivity analysis does not show that results are little changed under a range of plausible assumptions about the relationship between missingness in the outcome and its true value. The evidence indicates that findings are likely to be robust (MI/ LOCF) but evidence is not considered sufficient to eliminate all risk of bias due to missing data. There are some concerns
	Risk of bias judgement	Some concerns	equal proportions of missing outcome data between groups; reasons for refusal/non-attendance of post-assessments are unknown, so it is possible that missingness depends on its true value; multiple imputation was done on the basis of change from baseline to non-missing interim and post-tests, taking into account correlations between non-missing values for primary and all secondary outcome scores at baseline, interim post-tests and final post-test data. This approach is likely to lead to a lower risk of bias compared to MI based only on intervention group or compared to the LOCF method. Results were compared to results from LOCF procedure; similar results regarding significance of between-group differences. All in all, there are concerns.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The CAPS questionnaire was administered at baseline and at 3-months posttest. The research assistants administering the CAPS received training with an expert in CAPS administration

			(PH). The CAPS interview Cronbach's α ranges from 0.87 to 0.94, which indicates adequate internal consistency."
4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN		Quote: " Administration of the CAPS interview was supervised by two psychologists (TR and PH, study pyschologists)." comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
4.3 Were outcome assessors aware of the intervention received by study participants?	PY		Quote: "All baseline and post-test data were collected by two research assistants (MG and AR in the acknowledgments), masked to treatment assignment and uninvolved in any aspect of treatment delivery. Additionally, all participants were asked not to divulge their treatment assignment." comment: although the assessors were blind to the participants' condition, there was no blind assessment, since the participants (answering interview questions) were probably aware of their treatment allocation
4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY		comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention (with same number of sessions administered). Thus, it is less likely that participants might have answered according to their beliefs/expectations about the intervention effect.
4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN		
Risk of bias judgement	**Some concerns**		although the interviewers were blinded, assessment was not completely blind (participants might have been aware of their treatment allocation), increases the risk that participants answered interview questions according to their expectations regarding the treatment efficacy or according to their beliefs about desired results (to please the investigators). The risk of bias due to lack of

			participant blinding, however, is significantly lowered by the fact that the comparator was also an active treatment condition.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	Quote: "This study is registered with ClinicalTrials.gov, number NCT01865123." "Rutledge T, Nidich S, Schneider R, et al. Design and rationale of a comparative effectiveness trial evaluating transcendental meditation against established therapies for PTSD. Contemp Clin Trials 2014; 39: 50–56." comment: The trial was pre-registered (clinicaltrials.gov). examination of the history of changes and the comparison of the report with the published study protocol (published 2014, investigators were recruiting between 2013-2016) indicates that authors adhered to their pre-specified intentions in all aspects that are presented and that are relevant for the result of interest
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	comment: see above, all details reported regarding measurements are consistent with the trial protocol published before analysis
	5.3 ... multiple eligible analyses of the data?	N	comment: see above, all details reported regarding data analysis producing the result of interest are consistent with the trial protocol (published before analysis)
	Risk of bias judgement	**Low**	The trial was prospectively registered (clinicaltrials.gov) and a published study protocol is available. examination of the history of changes and comparison of the report with the published study protocol indicates that authors adhered to their pre-specified intentions in all aspects that are presented and that are relevant for the result of interest

Overall bias	Risk of bias judgement	Some concerns	equal proportions of missing outcome data between groups, reasons for refusal/non-attendance of post-assessments are unknown, ITT analyses using multiple imputation (taking several variables/correlations into account) and LOCF methods yielded similar results; some concerns remain. risk of bias due to lack of participant blinding is lowered by the fact that the comparator was also an active treatment condition, but there are concerns. The trial was pre-registered (clinicaltrials.gov) and a published study protocol is available; reported information is consistent with their pre-specified intentions, therefore risk of bias due to selective reporting is low.

Unique ID	96	Study ID	1070103	Assessor	R
Ref or Label	Nidich 2018	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	MBI	Comparator	Psychoeducation	Source	Journal article(s) with results of the trial; Trial protocol; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "Stratified block randomisation, stratified on gender and years since military service release, was used to assign participants to their study treatment. Treatment groups were matched on number of treatment visits (12 sessions), length of each session (90 minutes), and duration of treatment (12 weeks)."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	"Allocation concealment was achieved by an off-site coinvestigator (JS), who randomly assigned each participant to a

			treatment group and informed the study coordinator (EM) of the treatment assignments (who then notified the participant)."
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	Quote: "Groups were compared on baseline and demographic variables, with ANOVA for continuous variables and χ² tests for categorical variables." comment: results of significance test not reported, but reported baseline characteristics (table 1) reveal no major differences between groups
	Risk of bias judgement	**Low**	stratified block randomization; allocation concealment; no substantial differences at baseline
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	Quote: "This single-blind, three-arm randomised controlled trial [...]"
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	"Additionally, all participants were asked not to divulge their treatment assignment." comment: blinding of participants and therapists is not feasible in the context of the intervention study
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "All participants who were randomly assigned were included in the analyses, following the intention-to-treat principle. We did multiple imputations to include missing values at each time point, including the interim post-test visits, using the SAS software

			MI Procedure (Markov Chain Monte Carlo method; [SAS version 9.1.3])."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	blinding of participants and therapists is not feasible; ITT analyses conducted
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "166 (81%) participants completed the final 3-month posttest (53 (78%) of 68 for TM, 57 (84%) of 68 for PE, 56 (85%) of 66 for HE). All 202 eligible patients randomly assigned to treatment were included in the intention-totreat analyses regardless of treatment dropout or missing post-test data."

	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	Quote: " However, we also used the LOCF imputation method to estimate treatment effects in a sensitivity analysis because multiple imputation might yield less conservative estimates [...]" "A secondary analysis of change based on the LOCF method of imputing data yielded slightly more conservative results in terms of reductions in PCL-M and PHQ-9 scores compared with the intention-to-treat analysis based on multiple imputation; however, the LOCF method yielded similar results regarding statistical significance of between-group differences (appendix)." "Missing final posttest scores were imputed on the basis of change from baseline to non-missing interim and post-tests, taking into account correlations between non-missing values for primary and all secondary outcome scores at baseline, interim post-tests (PCL-M and PHQ-9 only) and final post-test." comment: Imputing the outcome variable through methods such as 'last-observation-carried-forward' or via multiple imputation based only on intervention group, should not be assumed to correct for bias due to missing outcome data. In this case MI was done taking additional variables/correlations into account, reducing the risk of bias. However, no sensitivity analyses was done showing that results are little changed under a range of plausible assumptions about the relationship between missingness in the outcome and its true value. The evidence is not considered sufficient to eliminate a risk of bias
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	Quote: "There were no treatment-related adverse events for any of the three treatments. The numbers of serious adverse events reported during the trial were not significantly different among treatment groups. There were three serious adverse events in the TM group (two suicide attempts, one death [non-suicidal]), two in the PE group (one drug overdose, one illness), and two in the HE
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	

			group (two psychiatric hospitalisations)."
			comment: reasons for declined/missed post-assessments of participants are unknown or not reported in detail for most participants with missing posttest data which increases the risk of bias. Reasons may be related to the treatment or the participants' health status. Proportions of missing posttest data did not differ between groups and there were no differences in adverse events. For details regarding the analysis see above. Multiple Imputation was done taking additional variables/correlations into account, reducing the risk of bias due to missing outcome data. However, their sensitivity analysis does not show that results are little changed under a range of plausible assumptions about the relationship between missingness in the outcome and its true value. The evidence indicates that findings are likely to be robust (MI/ LOCF) but evidence is not considered sufficient to eliminate all risk of bias due to missing data. There are some concerns
	Risk of bias judgement	**Some concerns**	equal proportions of missing outcome data between groups; reasons for refusal/non-attendance of post-assessments are unknown, so it is possible that missingness depends on its true value; multiple imputation was done on the basis of change from baseline to non-missing interim and post-tests, taking into account correlations between non-missing values for primary and all secondary outcome scores at baseline, interim post-tests and final post-test data. This approach is likely to lead to a lower risk of bias compared to MI based only on intervention group or compared to the LOCF method. Results were compared to results from LOCF procedure; similar results regarding significance of between-group differences. All in all, there are concerns.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The CAPS questionnaire was administered at baseline and at 3-months posttest. The research assistants administering the CAPS received training with an expert in CAPS administration

			(PH). The CAPS interview Cronbach's α ranges from 0.87 to 0.94, which indicates adequate internal consistency."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: " Administration of the CAPS interview was supervised by two psychologists (TR and PH, study pyschologists)." comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "All baseline and post-test data were collected by two research assistants (MG and AR in the acknowledgments), masked to treatment assignment and uninvolved in any aspect of treatment delivery. Additionally, all participants were asked not to divulge their treatment assignment." comment: although the assessors were blind to the participants' condition, there was no blind assessment, since the participants (answering interview questions) were probably aware of their treatment allocation
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention (with same number of sessions administered). Thus, it is less likely that participants might have answered according to their beliefs/expectations about the intervention effect.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN	
	Risk of bias judgement	**Some concerns**	although the interviewers were blinded, assessment was not completely blind (participants might have been aware of their treatment allocation), increases the risk that participants answered interview questions according to their expectations regarding the treatment efficacy or according to their beliefs about desired results (to please the investigators). The risk of bias due to lack of

			participant blinding, however, is significantly lowered by the fact that the comparator was also an active treatment condition.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	Y	Quote: "This study is registered with ClinicalTrials.gov, number NCT01865123." "Rutledge T, Nidich S, Schneider R, et al. Design and rationale of a comparative effectiveness trial evaluating transcendental meditation against established therapies for PTSD. Contemp Clin Trials 2014; 39: 50–56." comment: The trial was pre-registered (clinicaltrials.gov). examination of the history of changes and the comparison of the report with the published study protocol (published 2014, investigators were recruiting between 2013-2016) indicates that authors adhered to their pre-specified intentions in all aspects that are presented and that are relevant for the result of interest
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	comment: see above, all details reported regarding measurements are consistent with the trial protocol published before analysis
	5.3 ... multiple eligible analyses of the data?	N	comment: see above, all details reported regarding data analysis producing the result of interest are consistent with the trial protocol (published before analysis)
	Risk of bias judgement	**Low**	The trial was prospectively registered (clinicaltrials.gov) and a published study protocol is available. examination of the history of changes and comparison of the report with the published study protocol indicates that authors adhered to their pre-specified intentions in all aspects that are presented and that are relevant for the result of interest

Overall bias	Risk of bias judgement	Some concerns	equal proportions of missing outcome data between groups, reasons for refusal/non-attendance of post-assessments are unknown, ITT analyses using multiple imputation (taking several variables/correlations into account) and LOCF methods yielded similar results; some concerns remain. risk of bias due to lack of participant blinding is lowered by the fact that the comparator was also an active treatment condition, but there are concerns. The trial was pre-registered (clinicaltrials.gov) and a published study protocol is available; reported information is consistent with their pre-specified intentions, therefore risk of bias due to selective reporting is low.

Unique ID	100	Study ID	1100101	Assessor	R
Ref or Label	Pacella 2012	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	PE	Comparator	WL	Source	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: " The principal investigator (DLD) generated the allocation sequence using blocked randomization (4:3 ratio of experimental: control participants) [...]"
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: " Results from chi square analyses revealed that the PE and control groups did not differ by gender or race. Results from one-way ANOVAs also revealed no significant differences between groups on age, time living with HIV, sexual orientation,

			HIV-related PTSS, nonHIV-related PTSS, depression, posttraumatic cognitions, and substance use (see Table 1)."
	Risk of bias judgement	**Some concerns**	blocked randomization, unclear allocation concealment; no baseline differences.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as blinding is not feasible in this context of administration of psychological interventions
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	An intent-to-treat (ITT) method was used in order to remove potential bias due to participant attrition and non-compliance with the study protocol [76]. The conservative approach of last observation carried forward (LOCF) was applied [...]"
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	probably no blinding of participants or therapists; ITT analysis.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: " At the post-intervention assessment, 23 participants were retained in the PE group (32% drop-out rate) and 24 participants were retained in the control group (0% drop-out rate)."
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis reported
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	

	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PY	Quote: "Unequal numbers of participants were assigned to each group, as it was anticipated that the PE group would have a higher dropout rate." "At the post-intervention follow-up, a greater number of participants in the PE group dropped out of the study than participants in the control group ($\chi2$ (1, n = 65) = 11.38, P < .001). Participants who dropped out of the study reported lower levels of PTSS in reference to the nonHIV-related trauma than participants who were retained at this time point (F(1,62) = 5.49, P = .02). There were no significant differences in any other demographic or study variables between participants who dropped out and those who were retained at any time point. Further, drop-out rates did not differ between participants whose most distressing trauma was HIV-related compared to nonHIV-related." comment: missingness in the outcome could depend on its true value, as some participants dropped out for reasons related to group allocation (n=5 who dropped out because "therapy was too complex/lost interest in participating" [flow-chart]). In addition, significantly more participants dropped out from PE (32%) than from WL (0%). Therefore, the risk of bias is rated "high".
	Risk of bias judgement	**High**	unequal proportions of dropout; reasons for dropout related to the assigned treatment condition in some cases; no sensitivity analysis or analysis correcting for bias.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "At each assessment participants completed the PTSD Symptom Scale-Interview (PSS-I [70]). [...] Internal consistency for the PSS-I was acceptable in our sample (HIVrelated trauma baseline alpha = .81; Post-intervention = .89; 3-month = .86; 6-months = .84; NonHIV-related trauma baseline alpha = .81; Post-intervention = .93; 3-month = .86; 6months = .78."

4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: " PE participants completed the self-report PSS at every other therapy session (i.e., 2, 4, 6, 8, and 10) to track symptom progression through therapy, and reported on substance use at the start of each therapy session." comment: PE participants were assessed more often; apart from that, the same measurement methods and thresholds were used for all participants and it is unlikely that results were affected by measurement error
4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "All follow-up assessments were conducted at the social service agency by the same blind interviewer who conducted the baseline assessment." comment: although the interviewer was blind to the participants' condition, there was no blind assessment, since the participants (answering interview questions) are the actual assessor and they were probably aware of their treatment allocation
4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: assessment could have been influenced because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect
4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is high, as the comparator is a passive control condition (here: WL).
Risk of bias judgement	**High**	despite of the blinding of interviewers there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. The risk is particularly relevant as the comparator was a no-treatment waitlist condition.

Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	
	5.3 ... multiple eligible analyses of the data?	PN	comment: both ITT and completers data reported; generally, the risk of bias due to multiple eligible analyses of the data is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure of interest; for all time points of interest; results for the ITT sample as well as for completers are reported; generally, the risk of bias due selection of results based on multiple eligible analyses is lowered by the fact that the results assessed here are raw values (means, SDs). All in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are some concerns.
Overall bias	**Risk of bias judgement**	**High**	no information on allocation concealment; risk of bias due to participants' knowledge of intervention; potential bias due to missing outcome data as reasons for dropout may be related to treatment and proportions of dropout were unequal; not enough information regarding pre-specified analysis plan to reliably assess the risk of bias due to selection of the reported result.

Unique ID	104	**Study ID**	1120101	**Assessor**	R
Ref or Label	Pearson 2019	**Aim**	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	CPT	**Comparator**	WL	**Source**	Journal article(s) with results of the trial; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)

Outcome	self-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "Random assignment to condition was computer-generated with block sizes ranging from 2 to 8 to reduce the detection of a pattern and prepared by an external statistician. Allocation concealment involved the use of sequentially numbered, opaque, sealed envelopes containing the group assignment, which the research manager opened at the moment of randomization"
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "Sociodemographic and other descriptive information for the final analytic sample of 73 female participants by study condition are provided in Table 1" comment: no major differences
	Risk of bias judgement	**Low**	computer-generated with block randomization; allocation concealment using sealed envelopes; no substantial baseline differences between groups.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	Y	Quote: "Due to the nature of the intervention, participants and the study team could not be blinded to condition."
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	

	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "[...] longitudinal regression analyses were conducted using generalized estimating equations (GEE) [52] with cluster robust standard errors [53]. All participants randomized at baseline were included in the primary outcome analyses (i.e., an intent-to-treat approach) using all available data, including participants who missed one or more assessments."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding of participants or personnel (not feasable); mITT analysis.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "With respect to the primary study outcomes, 21% of participants (n = 15) had complete data from all assessments (immediate intervention: 3 assessments; waitlist control: 4 assessments), 42% (n = 31) were missing one assessment, and 37% were missing multiple assessments (n = 27)." comment: see Figure 1: of n=37 allocated to CPT, n=6 [16.21%] were post-assessed; of n=36 alssigned to WL, n=34 [94.44%] were post-assessed. Overall, percentage of missing pre-test data = 0%; missing posttest data = 45.2%
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis reported
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	Quote: " No serious adverse events were identified."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PY	" To assess for differences between participants who completed all assessments and those who were missing a single or multiple assessments, Pearson χ2 tests and one-way analyses of variance were conducted, respectively, on categorical and continuous socio-demographic characteristics and baseline levels of the outcomes. There were no statistically significantly differences between the degree of missing data and any demographic

			characteristic or baseline outcome." comment: see Figure 1: of n=37 allocated to CPT, n=6 [16.21%] were post-assessed; of n=36 alssigned to WL, n=34 [94.44%] were post-assessed. The difference in proportions of dropout between both groups suggests that missingness in the outcome might depend on its true value. In addition, no reasons of dropout are reported. Analysis of differences between participants with and without missing posttest data did not include clinical variables, but only sociodemographic characteristics. All in all, there is a high risk that missingness in the outcoe could depend on its true value.
	Risk of bias judgement	**High**	high attrition in the CPT group; no analysis correcting for bias or sensitivity analysis; reasons for dropout not reported, difference in proportions of missing posttest data between groups.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	PN	Quote: "Most items and all scales were selected from published and validated measures, and all items were pretested for cultural appropriateness in this community." "The 17-item PTSD Symptom Scale Self-Report Version [47] was used to assess the presence and severity of past month PTSD symptoms based on DSM-IV criteria. [...] with higher scores reflecting greater severity of PTSD symptoms (published α = 0.85, study α = 0.91)"
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "All participants completed 45-min audio computeradministered assessments at baseline, immediate post-, and 3-month post-intervention." comment: same measurement methods and thresholds were used for all participants, and used at comparable time points

	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	Quote: "Due to the nature of the intervention, participants and the study team could not be blinded to condition."
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: assessment could have been influenced because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is high, as the comparator is a passive control condition
	Risk of bias judgement	**High**	there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. It is particularly high as the comparator was a no-treatment condition (WL)
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: "use disorders we chose a conservative range of 3 months. The study protocol was approved by the University of Washington Institutional Review Board (#43091, NCT01849029) and a tribal review board." comment: Trial Registration. The researchers' pre-specified intentions are available in sufficient detail to believe that outcome measurements and analyses (of means, SDs) can be compared with those published in the report
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: results are reported for all participants randomized (ITT sample); generally, the risk of bias due selection of results based on multiple eligible analyses is lowered by the fact that results assessed here are raw values (means, SDs)

	Risk of bias judgement	Low	trial was prospectively registered (clinicaltrials.gov); examination of the history of changes and the comparison of study record versions indicates that authors adhered to their pre-specified intentions in all aspects that are presented and that are relevant for the result of interest.
Overall bias	Risk of bias judgement	High	high attrition in the CPT group, no analysis correcting for bias or sensitivity analysis, reasons for dropout not reported, difference in proportions of missing posttest data between groups; leading to a high risk of bias due to missing outcome data. risk of bias due to knowledge of the intervention because participants were not blind to their condition; passive control condition increasing this risk. All in all, there might be bias favouring the experimental condition.

Unique ID	114	Study ID	1210101	Assessor	R
Ref or Label	Resick 2002	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	CPT	Comparator	PE	Source	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	NI	Quote: "At the end of the second assessment, the MA participants were randomly assigned to either PE or CPT."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	comment: The randomization procedure is described incompletely, without confirming that there was a random

			component. A simple statement such as "we randomly allocated" is considered insufficient to be confident that the allocation sequence was genuinely randomized.
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	Quote: "In the ITT sample, there were no significant differences in demographic characteristics among the three groups" comment: visual inspection of pre-treatment scores on clinical measures indicates that there were no substantial baseline differences
	Risk of bias judgement	**Some concerns**	insufficient information regarding random sequence generation; no information on allocation concealment; no substantial baseline differences on any demographic or clinical variables
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants may have been aware of the assigned condition as adequate blinding is not feasible when psychological interventions are implemented
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias; investigators made vigorous efforts to assess and prevent deviations from intended interventions which indicates that there is a low risk of bias due to deviations
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "The results were analyzed in three different ways for comparison purposes. Unfortunately, this study was designed and conducted before ITT analysis became standard. Therefore, we did not continue to assess women who dropped out of treatment [...] Initially, all of the participants who were accepted and randomized into the trial were analyzed with their last observations carried forward (LOCF)."

	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding of therapists or participants; no deviations from the intended intervention that arose because of the trial context; ITT analysis with LOCF method used.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: " Of 181 women randomized into the trial, 10 were terminated from the study as a result of meeting exclusion criteria subsequent to new violence (women had to be at least 3 months posttrauma), changes in medication, or substance dependence relapse. Therefore, the intent-to-treat (ITT) sample included 171 women, among whom 13 never returned for the first session. Thirty-seven women dropped out of treatment, and 121 women completed treatment along with at least the posttreatment assessment: 41 CPT clients, 40 PE clients, and 40 MA clients."comment: amount of missing data in the CPT group = 33.87% [n=41 participants ot of n=62 randomized]; and in the PE group = 35.48% [n=40 participants ot of n=62 randomized];
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	Quote: " Another method of handling nonrandom missing data due to dropout is to use mixedeffects linear regression analysis or random regression" "Supplementing the use of LOCF data with random regression models as a converging test of our hypotheses allowed us added protection against misleading findings" "Finally, those women who completed treatment were analyzed separately" "A consistent picture emerged through the use of different types of statistical analyses with ITT and completer samples" " Although it was beyond the scope of this study to examine the full range of variables that might have affected treatment completion, this would be an important topic for future research."

			comment: no sensitivity analyses showing that results are little changed under a range of plausible assumptions about the relationship between missingness in the outcome and its true value; In addition, arguably, in the presence of non-random dropout, a wholly satisfactory analysis of the data is not feasible.
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: "Therefore, the intent-to-treat (ITT) sample included 171 women, among whom 13 never returned for the first session" "it was beyond the scope of this study to examine the full range of variables that might have affected treatment completion" "there were no differences in pretreatment PTSD or depression between those who dropped out and those who completed treatment."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	"First, the four expectancy questions at pretreatment were subjected to a multivariate analysis of variance (MANOVA) to determine whether the women who dropped out had different expectations from those who completed treatment. The MANOVA was nonsignificant. Next, we conducted a repeated measures MANOVA (pretreatment– posttreatment) with type of therapy (CPT or PE) as the independent variable. There was no interaction between groups and sessions"

			equal proportions of dropout between both groups; overall there is a considerable amount of missing data; no information regarding group assignment of participants who never started treatment [n=13]; As the quoted information indicates, treatment expectations did not significantly differ between dropouts and completers; however, it was not tested whether participants with and wothout missing data differed on an demographic variables or on clinical variables other than pretreatment PTSD/depression. reasons for dropout are not reported, so it is unknown whether (1) reasons for dropout differed between groups and (2) whether reasons were study-related.
	Risk of bias judgement	**High**	considerable amount of missing data [CPT= 33.87%; PE= 35.48%], no evidence that the rsult was not biased; insufficient information regarding reasons for dropout or other sources of bias.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The CAPS (Blake et al., 1990) is an interviewer-administered diagnostic instrument that measures PTSD. It has been found to have excellent psychometric properties (Blake et al., 1995)"
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: " After reliability had been established (100% diagnostic reliability and high item reliability), all diagnostic interviewers had audiotapes reviewed by senior project staff on a random, ongoing basis to ensure that there was no drift in diagnostic decisions. [...] A random sample of 66 tapes was selected for evaluation of interrater reliability for the CAPS. Categorical diagnostic analyses revealed that the kappa coefficient for the overall PTSD diagnosis was .74, with 92% interrater agreement" comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points

			Quote: "The MA condition served as a waiting-list control. Women who were assigned to this condition were told that therapy would be provided in 6 weeks and that an interviewer would call them every 2 weeks to ensure that they did not need emergency services."
	4.3 Were outcome assessors aware of the intervention received by study participants?	NI	comment: participants may have been aware of the assigned condition as adequate blinding is not feasible when psychological interventions are implemented; interviewers were potentially unblinded since (1) there is no mention of blinding of interviewers in the report and (2) the quote indicates that intwerwievers were aware when participants were assigned to the WL condition (so it can be assumed that there was no blinding in general). Regardless of the blinding of interviewers, however, there was no blind assessment as participants were not blinded.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention. However, the interviewer might have known the participants' treatment condition which raises serious concerns: For subjective outcomes such as 'clinical impression of improvement', knowledge of the intervention received could be highly influential.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NI	
	Risk of bias judgement	**High**	no blinding of participants; unclear whether interviewers were blinded; as knowledge of the intervention received could be highly influential for subjective outcomes such as 'clinical impression of improvement', the risk of bias is rated 'high'.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	comment: no trial registration and insufficient information regarding pre-specified analysis plan
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	

	5.3 ... multiple eligible analyses of the data?	PN	comment: both ITT and completers data reported; CAPS results from ITT analysis using (1) the LOCF method as well as (2) random regression reported.
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure of interest; for all time points of interest; completers' results reported as well as results from both types of ITT analyses; All in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are some concerns.
Overall bias	**Risk of bias judgement**	**High**	insufficient information regarding random sequence generation; no information on allocation concealment; no substantial baseline differences on any demographic or clinical variables. no blinding of therapists or participants; no deviations from the intended intervention that arose because of the trial context; ITT analysis with LOCF method used. considerable amount of missing data [CPT= 33.87%; PE= 35.48%], no evidence that the rsult was not biased; insufficient information regarding reasons for dropout or other sources of bias. no blinding of participants; unclear whether interviewers were blinded; as knowledge of the intervention received could be highly influential for subjective outcomes such as 'clinical impression of improvement', the risk of bias is rated 'high'. not enough information regarding pre-specified analysis plan to reliably assess the risk of bias due to selection of the reported result.

Unique ID	115	**Study ID**	1210102	**Assessor**	R
Ref or Label	Resick 2002	**Aim**	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	CPT	**Comparator**	WL	**Source**	Journal article(s) with results of the trial

Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	NI	Quote: "At the end of the second assessment, the MA participants were randomly assigned to either PE or CPT."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	comment: The randomization procedure is described incompletely, without confirming that there was a random component. A simple statement such as "we randomly allocated" is considered insufficient to be confident that the allocation sequence was genuinely randomized.
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	Quote: "In the ITT sample, there were no significant differences in demographic characteristics among the three groups" comment: visual inspection of pre-treatment scores on clinical measures indicates that there were no substantial baseline differences
	Risk of bias judgement	Some concerns	insufficient information regarding random sequence generation; no information on allocation concealment; no substantial baseline differences on any demographic or clinical variables
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably aware of the assigned condition as adequate blinding is not feasible when psychological interventions are implemented, especially given the differences between active treatment and passive control condition
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias; investigators made vigorous efforts to assess and prevent deviations from

			intended interventions which indicates that there is a low risk of bias due to deviations
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "The results were analyzed in three different ways for comparison purposes. Unfortunately, this study was designed and conducted before ITT analysis became standard. Therefore, we did not continue to assess women who dropped out of treatment [...] Initially, all of the participants who were accepted and randomized into the trial were analyzed with their last observations carried forward (LOCF)."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding of therapists or participants; no deviations from the intended intervention that arose because of the trial context; ITT analysis with LOCF method used.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: " Of 181 women randomized into the trial, 10 were terminated from the study as a result of meeting exclusion criteria subsequent to new violence (women had to be at least 3 months posttrauma), changes in medication, or substance dependence relapse. Therefore, the intent-to-treat (ITT) sample included 171 women, among whom 13 never returned for the first session. Thirty-seven women dropped out of treatment, and 121 women completed treatment along with at least the posttreatment assessment: 41 CPT clients, 40 PE clients, and 40 MA clients." comment: amount of missing data in the CPT group = 33.87%

			[n=41 participants ot of n=62 randomized]; and in the WL group= 14.89% [n=40 out of n=47 randomized]
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	Quote: " Another method of handling nonrandom missing data due to dropout is to use mixedeffects linear regression analysis or random regression" "Supplementing the use of LOCF data with random regression models as a converging test of our hypotheses allowed us added protection against misleading findings" "Finally, those women who completed treatment were analyzed separately" "A consistent picture emerged through the use of different types of statistical analyses with ITT and completer samples" " Although it was beyond the scope of this study to examine the full range of variables that might have affected treatment completion, this would be an important topic for future research." comment: no sensitivity analyses showing that results are little changed under a range of plausible assumptions about the relationship between missingness in the outcome and its true value; In addition, arguably, in the presence of non-random dropout, a wholly satisfactory analysis of the data is not feasible.

	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: "Therefore, the intent-to-treat (ITT) sample included 171 women, among whom 13 never returned for the first session"
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	"it was beyond the scope of this study to examine the full range of variables that might have affected treatment completion" "there were no differences in pretreatment PTSD or depression between those who dropped out and those who completed treatment." "First, the four expectancy questions at pretreatment were subjected to a multivariate analysis of variance (MANOVA) to determine whether the women who dropped out had different expectations from those who completed treatment. The MANOVA was nonsignificant. Next, we conducted a repeated measures MANOVA (pretreatment– posttreatment) with type of therapy (CPT or PE) as the independent variable. There was no interaction between groups and sessions" unequal proportions of dropout between both groups [33.87% vs. 14.89%; no significance test]; considerable amount of missing data in the CPT group; no information regarding group assignment of participants who never started treatment [n=13]; As the quoted information indicates, treatment expectations did not significantly differ between dropouts and completers; however, it was not tested whether participants with and without missing data differed on an demographic variables or on clinical variables other than pretreatment PTSD/depression. reasons for dropout are not reported, so it is unknown whether (1) reasons for dropout differed between both groups and (2) whether reasons were study-related.
	Risk of bias judgement	**High**	amount of missing data: CPT= 33.87%; WL= 14.89%, no evidence that the result was not biased; insufficient information regarding reasons for dropout and other potential sources of bias.

Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The CAPS (Blake et al., 1990) is an interviewer-administered diagnostic instrument that measures PTSD. It has been found to have excellent psychometric properties (Blake et al., 1995)"
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: " After reliability had been established (100% diagnostic reliability and high item reliability), all diagnostic interviewers had audiotapes reviewed by senior project staff on a random, ongoing basis to ensure that there was no drift in diagnostic decisions. [...] A random sample of 66 tapes was selected for evaluation of interrater reliability for the CAPS. Categorical diagnostic analyses revealed that the kappa coefficient for the overall PTSD diagnosis was .74, with 92% interrater agreement" comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	NI	Quote: "The MA condition served as a waiting-list control. Women who were assigned to this condition were told that therapy would be provided in 6 weeks and that an interviewer would call them every 2 weeks to ensure that they did not need emergency services." comment: participants were probably aware of the assigned condition as adequate blinding is not feasible when psychological interventions are implemented, especially given the differences between active treatment and passive control condition- and as indicated by the quote above
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. It is particularly high as the comparator was a no-treatment condition (WL). In addition, the interviewer might
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NI	

			have known the participants' treatment condition which raises serious concerns: For subjective outcomes such as 'clinical impression of improvement', knowledge of the intervention received could be highly influential.
	Risk of bias judgement	**High**	no blinding of participants; unclear whether interviewers were blinded; risk of bias due to knowledge of the intervention is particularly high as the comparator was a no-treatment condition; in addition, knowledge of the intervention received could be highly influential for subjective outcomes such as 'clinical impression of improvement'.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	comment: no trial registration and insufficient information regarding pre-specified analysis plan
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	
	5.3 ... multiple eligible analyses of the data?	PN	comment: both ITT and completers data reported; CAPS results from ITT analysis using (1) the LOCF method as well as (2) random regression reported.
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure of interest; for all time points of interest; completers' results reported as well as results from both types of ITT analyses; All in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are some concerns.

Overall bias	Risk of bias judgement	High	insufficient information regarding random sequence generation; no information on allocation concealment; no substantial baseline differences on any demographic or clinical variables. no blinding of therapists or participants; no deviations from the intended intervention that arose because of the trial context; ITT analysis with LOCF method used. amount of missing data: CPT= 33.87%; WL= 14.89%, no evidence that the result was not biased; insufficient information regarding reasons for dropout and other potential sources of bias. no blinding of participants; unclear whether interviewers were blinded; risk of bias due to knowledge of the intervention is particularly high as the comparator was a no-treatment condition; in addition, knowledge of the intervention received could be highly influential for subjective outcomes such as 'clinical impression of improvement'. not enough information regarding pre-specified analysis plan to reliably assess the risk of bias due to selection of the reported result.

Unique ID	116	Study ID	1210103	Assessor	R
Ref or Label	Resick 2002	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	PE	Comparator	WL	Source	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the	1.1 Was the allocation sequence random?	NI	Quote: "At the end of the second assessment, the MA participants were randomly assigned to either PE or CPT."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	

randomization process			comment: The randomization procedure is described incompletely, without confirming that there was a random component. A simple statement such as "we randomly allocated" is considered insufficient to be confident that the allocation sequence was genuinely randomized.
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	Quote: "In the ITT sample, there were no significant differences in demographic characteristics among the three groups" comment: visual inspection of pre-treatment scores on clinical measures indicates that there were no substantial baseline differences
	Risk of bias judgement	**Some concerns**	insufficient information regarding random sequence generation; no information on allocation concealment; no substantial baseline differences on any demographic or clinical variables
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably aware of the assigned condition as adequate blinding is not feasible when psychological interventions are implemented, especially given the differences between active treatment and passive control condition
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias; investigators made vigorous efforts to assess and prevent deviations from intended interventions which indicates that there is a low risk of bias due to deviations
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	

	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "The results were analyzed in three different ways for comparison purposes. Unfortunately, this study was designed and conducted before ITT analysis became standard. Therefore, we did not continue to assess women who dropped out of treatment [...] Initially, all of the participants who were accepted and randomized into the trial were analyzed with their last observations carried forward (LOCF)."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding of therapists or participants; no deviations from the intended intervention that arose because of the trial context; ITT analysis with LOCF method used.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: " Of 181 women randomized into the trial, 10 were terminated from the study as a result of meeting exclusion criteria subsequent to new violence (women had to be at least 3 months posttrauma), changes in medication, or substance dependence relapse. Therefore, the intent-to-treat (ITT) sample included 171 women, among whom 13 never returned for the first session. Thirty-seven women dropped out of treatment, and 121 women completed treatment along with at least the posttreatment assessment: 41 CPT clients, 40 PE clients, and 40 MA clients." comment: amount of missing data in the CPT group = 35.48% [n=40 participants ot of n=62 randomized]; and in the WL group= 14.89% [n=40 out of n=47 randomized]

	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	Quote: " Another method of handling nonrandom missing data due to dropout is to use mixedeffects linear regression analysis or random regression" "Supplementing the use of LOCF data with random regression models as a converging test of our hypotheses allowed us added protection against misleading findings" "Finally, those women who completed treatment were analyzed separately" "A consistent picture emerged through the use of different types of statistical analyses with ITT and completer samples" " Although it was beyond the scope of this study to examine the full range of variables that might have affected treatment completion, this would be an important topic for future research." comment: no sensitivity analyses showing that results are little changed under a range of plausible assumptions about the relationship between missingness in the outcome and its true value; In addition, arguably, in the presence of non-random dropout, a wholly satisfactory analysis of the data is not feasible.
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: "Therefore, the intent-to-treat (ITT) sample included 171 women, among whom 13 never returned for the first session"
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	"it was beyond the scope of this study to examine the full range of variables that might have affected treatment completion" "there were no differences in pretreatment PTSD or depression between those who dropped out and those who completed treatment." "First, the four expectancy questions at pretreatment were subjected to a multivariate analysis of variance (MANOVA) to

			determine whether the women who dropped out had different expectations from those who completed treatment. The MANOVA was nonsignificant. Next, we conducted a repeated measures MANOVA (pretreatment– posttreatment) with type of therapy (CPT or PE) as the independent variable. There was no interaction between groups and sessions"
			unequal proportions of dropout between both groups [35.48% vs. 14.89%; no significance test]; considerable amount of missing data in the CPT group; no information regarding group assignment of participants who never started treatment [n=13]; As the quoted information indicates, treatment expectations did not significantly differ between dropouts and completers; however, it was not tested whether participants with and without missing data differed on an demographic variables or on clinical variables other than pretreatment PTSD/depression. reasons for dropout are not reported, so it is unknown whether (1) reasons for dropout differed between both groups and (2) whether reasons were study-related.
	Risk of bias judgement	**High**	amount of missing data: CPT= 35.48%; WL= 14.89%, no evidence that the result was not biased; insufficient information regarding reasons for dropout and other potential sources of bias.
Bias in measurement	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The CAPS (Blake et al., 1990) is an interviewer-administered diagnostic instrument that measures PTSD. It has been found to have excellent psychometric properties (Blake et al., 1995)"

of the outcome	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: " After reliability had been established (100% diagnostic reliability and high item reliability), all diagnostic interviewers had audiotapes reviewed by senior project staff on a random, ongoing basis to ensure that there was no drift in diagnostic decisions. [...] A random sample of 66 tapes was selected for evaluation of interrater reliability for the CAPS. Categorical diagnostic analyses revealed that the kappa coefficient for the overall PTSD diagnosis was .74, with 92% interrater agreement" comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	NI	Quote: "The MA condition served as a waiting-list control. Women who were assigned to this condition were told that therapy would be provided in 6 weeks and that an interviewer would call them every 2 weeks to ensure that they did not need emergency services." comment: participants were probably aware of the assigned condition as adequate blinding is not feasible when psychological interventions are implemented, especially given the differences between active treatment and passive control condition- and as indicated by the quote above
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. It is particularly high as the comparator was a no-treatment condition (WL). In addition, the interviewer might have known the participants' treatment condition which raises serious concerns: For subjective outcomes such as 'clinical impression of improvement', knowledge of the intervention received could be highly influential.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NI	

	Risk of bias judgement	**High**	no blinding of participants; unclear whether interviewers were blinded; risk of bias due to knowledge of the intervention is particularly high as the comparator was a no-treatment condition; in addition, knowledge of the intervention received could be highly influential for subjective outcomes such as 'clinical impression of improvement'.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	comment: no trial registration and insufficient information regarding pre-specified analysis plan
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	
	5.3 ... multiple eligible analyses of the data?	PN	comment: both ITT and completers data reported; CAPS results from ITT analysis using (1) the LOCF method as well as (2) random regression reported.
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure of interest; for all time points of interest; completers' results reported as well as results from both types of ITT analyses; All in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are some concerns.
Overall bias	**Risk of bias judgement**	**High**	insufficient information regarding random sequence generation; no information on allocation concealment; no substantial baseline differences on any demographic or clinical variables. no blinding of therapists or participants; no deviations from the intended intervention that arose because of the trial context; ITT analysis with LOCF method used. amount of missing data: CPT= 35.48%; WL= 14.89%, no evidence that the result was not biased; insufficient information regarding reasons for dropout and other potential sources of bias. no blinding of participants; unclear whether interviewers were blinded; risk of bias due to knowledge of the intervention is particularly high as the comparator was a no-treatment condition; in addition, knowledge of the intervention received could be highly influential for subjective outcomes such as 'clinical impression of improvement'. not enough information

					regarding pre-specified analysis plan to reliably assess the risk of bias due to selection of the reported result.

Unique ID	120	Study ID	1230101	Assessor	R
Ref or Label	Resick 2008	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	CPT	Comparator	CT	Source	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	NI	Quote: "After assessments were conducted and participants were accepted into treatment, they were randomly assigned to one of the three treatments by the data manager (the investigators and assessors were blind as to assignment and assessors continued to be blind to condition throughout the trial)."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	

			comment: the reported information on the randomization process is not explicit and detailed enough to confirm that the allocation process included a random component and that (and how) allocation concealment was warranted.
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "At pretreatment there were no differences between the three treatment conditions on the total score on the Therapeutic Outcome Questionnaire" "There were no differences between groups on any of the symptom measures at pretreatment" "Randomization was largely successful with regard to demographics of the sample and symptoms at the pretreatment assessment. In the ITT sample, there were no significant differences in demographics among the three groups except for income. [...] with the CPT group having significantly lower income than the other two conditions"
	Risk of bias judgement	**Some concerns**	no detailed information regarding random sequence generation and allocation concealment- the statement indicates that this was probably the case, but some concerns remain; no substantial baseline differences on any demographic or clinical or other study measures.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	NI	comment: therapists were most likely aware and participants were potentially aware of the assigned condition as blinding is not feasible when psychological interventions are implemented. The similarities between both treatment conditions may have decreased the probability that participants were aware of their group status.
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	

2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	PN	Quote: " Of 162 women randomized into the trial, 12 were terminated from the study, by design, for meeting exclusion criteria subsequent to new violence (women had to be at least 3 months posttrauma), changes in medication, or psychosis. Among them, 1 WA participant was terminated from the trial when the therapist stopped the protocol because of increased suicidal ideation. These terminations were evenly distributed across groups. Therefore, the intent-to-treat (ITT) sample included 150 women." "The primary analyses of the study were conducted with the ITT" comment: the sample labled "ITT sample" by the authors does not fulfill the ITT principles as 7 participants who were eligible at randomization were excluded post-randomization. The reasons (e.g. change in medication status, psychosis) may have been influenced by group assignment. Analyses excluding eligible trial participants post-randomization should also be considered inappropriate. postrandomization exclusions of ineligible participants (when eligibility was not confirmed until after randomization, and could not have been influenced by intervention group assignment) can be considered appropriate but this is not the case here.	
2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	PN	The number of eligible participants excluded post-randomization is small (6.5%). There is no precise rule to assess this potential and there is limited information to assess whether exclusions were strongly related to prognostic factors. There are concerns.	
Risk of bias judgement	**Some concerns**	ITT analysis using LOCF method; 7 participants (6.5%) who were eligible at randomization were excluded post-randomization and are not included in the "ITT analysis"; it is unknown whether exclusions were strongly related to prognostic factors. There are concerns.	

Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	comment: see table 4: of the n=56 participants randomized to CPT n=53 were included in the ITT analysis and of those there was posttreatment data for n=42 participants [=25% of missing data in the original ITT sample; =20.75% of missing data in the 'ITT' sample the analysis was conducted on]; of the n=51 participants randomized to CT n=47 were included in the 'ITT' analysis and there is posttreatment data for n=37 of those CT participants [=27.45% of missing data in the original ITT sample; =21.27% of missing data in the 'ITT' sample the analysis was conducted on]
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	Quote: "Diagnostic interviews were compared by chi-square analyses with last observation carried forward for missing data at any of the follow-ups." comment: no analysis correcting for bias or sensitivity analysis reported
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: "Of the women in the ITT sample, 24 never returned for the first session of therapy, 40 received partial therapy, of whom, 5 women received partial treatment because the allocated 12-week therapy time limit expired. [...] There were no significant differences between treatment groups on these treatment status categories. There were, however, differences in demographics among treatment status groups. There was a significant race effect on treatment completion, $\chi2(4, N = 150) = 15.55$, $p = .004$. Only 37.3% (19 of 51) of the African American women completed all therapy sessions. [...] There were also differences in treatment completion based on household income." " Of 162 women randomized into the trial, 12 were terminated from the study, by design, for meeting exclusion criteria subsequent to new violence (women had to be at least 3 months
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	

			posttrauma), changes in medication, or psychosis. Among them, 1 WA participant was terminated from the trial when the therapist stopped the protocol because of increased suicidal ideation. These terminations were evenly distributed across groups. Therefore, the intent-to-treat (ITT) sample included 150 women. There was one other unrelated adverse event during the trial." comment: moderate amount of missing posttest data; equal proportions of missing posttest data between both groups; participants with and without missing posttest data differed on two variables (income, race), however, it should be noted that clinical variables were not included in these analyses (!); it is unknown whether participant without posttest data showed more severe symptoms at baseline; no information on reasons for dropout. All in all, it is likely that missingness in the outcome depended on its true value.
	Risk of bias judgement	**High**	participants with and without missing posttest data differed on two variables (income, race), clinical variables were not included in these analyses (!); reasons for dropout not reported; all in all, it is likely that missingness in the outcome depended on its true value.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: " CAPS diagnoses and symptom severity scores have demonstrated reliability and validity (Weathers, Keane, & Davidson, 2001). Cronbach's alpha on CAPS total score for this study was .91." comment: The administered scale (CAPS) is a validated, gold-standard PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points

	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "the investigators and assessors were blind as to assignment and assessors continued to be blind to condition throughout the trial" comment: participants were not blinded and potentially aware of the assigned condition; therefore, assessments were not completely blind
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention. Thus, the probability that participants have answered according to their beliefs/expectations about the intervention effect is significantly lower.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN	
	Risk of bias judgement	**Some concerns**	although the interviewers were blinded, there is still a risk of bias; Because the comparator is also an active control condition and similar in many aspects the risk that participants might have answered questions according to their beliefs/expectations regarding their assigned condition is considered low; nevertheless, it cannot be fully eliminated.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	NI	
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure of interest; for all time points of interest; results for the ITT sample (as defined by the authors) reported; All in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are some concerns.

Overall bias	Risk of bias judgement		High	no detailed information regarding random sequence generation and allocation concealment- the statement indicates that this was probably the case, but some concerns remain; no substantial baseline differences on any demographic or clinical or other study measures. ITT analysis using LOCF method; 7 participants (6.5%) who were eligible at randomization were excluded post-randomization and are not included in the "ITT analysis"; it is unknown whether exclusions were strongly related to prognostic factors. participants with and without missing posttest data differed on two variables (income, race), clinical variables were not included in these analyses (!); reasons for dropout not reported; all in all, it is likely that missingness in the outcome depended on its true value. interviewers blinded but participants not blinded; because the comparator is also an active control condition and similar in many aspects the risk that participants might have answered questions according to their beliefs/expectations regarding their assigned condition is considered low; nevertheless, there are concerns regarding bias due to knowledge of the intervention. not enough information regarding pre-specified analysis plan in order to reliably assess the risk of bias due to selection of the reported results.

Unique ID	130	Study ID	1390101	Assessor	R
Ref or Label	Schnurr 2007	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	PE	Comparator	PCT	Source	Journal article(s) with results of the trial; Trial protocol; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)

Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1	

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "Study staff called a computerized voice information system at the study coordinating center to obtain the treatment assignment for participants. The voice information system first verified entry criteria to ensure accuracy and reduce errors. Verified eligible participants were randomized within each site to prolonged exposure or present-centered therapy using permuted blocks with random block sizes of 4 or 6. All study data were stored at the study coordinating center."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	PY	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "Women randomized to prolonged exposure and present-centered therapy did not differ at baseline.""The percentage of proscribed elements was low and did not differ (0.5 vs 1.5, respectively; P=.33)."
	Risk of bias judgement	**Low**	randomization using permuted blocks with random block sizes of 4 or 6; data were stored at study coordinating center to ensure allocation concealment; no baseline differences between groups.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists were probably necessarily aware and participants might have been aware of the assigned condition as blinding is not feasible when psychological interventions are implemented
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	Quote: "Prolonged exposure and presentcentered therapy therapists did not differ in global ratings of competence or adherence, which averaged between very good and excellent."

comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias |
| | 2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome? | NA | |

	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "Primary analyses were performed on the intention-to-treat sample, using data from all randomized participants."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding of therapists or participants; no deviations that arose because of the trial context; ITT analysis using data from all participants that were randomized.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "141 Assigned to Receive Prolonged, [...] 21 Lost to Follow-up" "143 Assigned to Receive Present-Centered Therapy, [...] 17 Lost to Follow-up" comment: amount of missing posttest data: PE=14.89%; PCT=11.88%
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	comment: no analysis correcting for bias or sensitivity analysis. although the methods that were used by the authors to handle missing data aim at reducing the risk of bias they do not offer evidence that the results was not biased by missing outcome data. In addition, the methods used (Quote: "Multiple imputation [...] with the Markov chain Monte Carlo method was used to impute missing values." "Outcomes were analyzed using the generalized linear mixed model (SAS PROC MIXED with iteratively reweighted likelihoods GLIMMIX macro)" are based on strong assumptions. The missigness mechanism can be assumed to be non-ignorable (MNAR) based on the reported information.
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: "Treatment dropout was higher in prolonged exposure (n=53 [38%]) than in present-centered therapy (n=30 [21%]) (P=.002)."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	

			"There were 5 serious adverse events in prolonged exposure (4 psychiatric hospitalizations and 1 suicide attempt) and 14 in present-centered therapy (2 deaths [nonsuicidal], 9 psychiatric hospitalizations, and 3 suicide attempts). No events were regarded as study-related; the suicide attempt in prolonged exposure was coded as possibly related."
			comment: see figure 1. insufficient information regarding reasons for dropout; differences in reasons for dropout between the groups are possible; unknown whether participants with and without missing posttest data differed on any study variables; it should also be noted that the coding scheme (to assess whether an adverse event was study-related or not) is unknown and no details are reported as to why those 19 serious adverse events are cosidered to be "not study-related"; taken toghether, the available information indicates that there is a risk of bias due to missingness in the outcome.
	Risk of bias judgement	**High**	insufficient information regarding reasons for dropout; not reported whether participants with and without missing posttest data differed on any study variables.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: " The primary outcome measure was PTSD symptom severity on the CAPS structured interview." comment: The administered scale (CAPS) is a validated, gold-standard PTSD measure and likely to be sensitive to intervention effects

4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points Quote: "A master's- or doctoral-level assessor, blinded to treatment assignment, performed assessments before and after treatment and at 3- and 6-month follow-up appointments."
4.3 Were outcome assessors aware of the intervention received by study participants?	NI	Quote: "A master's- or doctoral-level assessor, blinded to treatment assignment, performed assessments before and after treatment and at 3- and 6-month follow-up appointments." "With these procedures, unblinding occurred for 33 patients in the prolonged exposure group and 17 in the present-centered therapy group. For 11 patients (12 interviews), interviews performed subsequent to the unblinding were also rated by the assessment monitor. Discrepancies between the moni tor and the assessor were small on average and did not differ between groups" Regardless of the blinding of interviewers, there was no blind assessment as participants were not blinded. Arguably, the fact that both interventions as similar in many aspects may have reduced the probability of participants guessing or knowing their group status but this remains speculative as there was no systematic assessment
4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: despite of the blinding of interviewers there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. The risk of bias due to lack of participant blinding, however, is lowered by the fact that the comparator was also an active
4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN	

			intervention (and, in addition, both interventions share common elements of psychotherapeutic treatment).
	Risk of bias judgement	**Some concerns**	no blind assessment as participants were not blinded; the risk of bias is lowered by the fact that the comparator was also an active intervention.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: "Trial Registration clinicaltrials.gov Identifier: NCT0003261" comment: published study protocol ("Schnurr PP, Friedman MJ, Engel CC, et al. Issues in the design of multisite clinical trials of psychotherapy: CSP#494 as an example. Contemp Clin Trials. 2005; 26: 626-636"). The trial was pre-registered (clinicaltrials.gov). examination of the history of changes in the trial registry and examination of the published study protocol indicate that authors adhered to their pre-specified intentions in aspects relevant for the result of interest
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: see above [5.1]. results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: see above [5.1]; both data from ITT analysis and from completers analysis reported;
	Risk of bias judgement	**Low**	The trial was prospectively registered at ClinicalTrials.gov and a study protocol was published two years before publication of this report. A comparison of this data with the reported information indicates that, all in all, researchers adhered to their pre-specified intentions. The risk of bias is therefore considered low.
Overall bias	**Risk of bias judgement**	**High**	insufficient information regarding reasons for dropout; not reported whether participants with and without missing posttest data differed on any study variables. no blind assessment as participants were not blinded; the risk of bias is lowered by the fact that the comparator was also an active intervention. The trial was pre-registered at ClinicalTrials.gov and a study protocol was published two years before publication of this report. A

comparison of this data with the reported information indicates that, all in all, researchers adhered to their pre-specified intentions. The risk of bias is therefore considered low.

Unique ID	132	Study ID	1400101	Assessor	R
Ref or Label	Seppälä 2014	**Aim**	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	MBI	**Comparator**	WL	**Source**	Journal article(s) with results of the trial
Outcome	self-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: " A biostatistician determined the simple randomization procedure using a computer-generated randomization list. The study coordinator then assigned eligible participants to the groups according to the randomization list: Sudarshan Kriya yoga active (n = 11) or waitlist control (n = 10) group."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "The absence of group differences at Time 1 for all measures (all ps > .108) indicated that the randomization procedure for group assignment was successful."
	Risk of bias judgement	Some concerns	computer-generated randomization list; insufficient information regarding allocation concealment; no baseline differences.
	2.1.Were participants aware of their assigned intervention during the trial?	PY	

Bias due to deviations from intended interventions	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as blinding is not feasible when psychological interventions are implemented; especially not in view of the apparent differences between an active and a passive treatment condition
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "To address the issue of missing and unusable data, we implemented an intent-to-treat analysis using the maximuml ikelihood estimation" comment: ITT analysis including all participants randomized, using linear mixed models
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding, neither of instructors nor participants; no deviations reported; ITT analysis including all participants randomized.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	Quote: "All 21 participants completed physiological and self-report assessments at Time 1. One participant in the active group dropped out after the third day because he disliked the intervention. Ten participants in each group completed assessments at Time 2"
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	

	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	n=21 were randomized, n=21 completed baseline assessment, n=20 completed posttest.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "For overall PTSD ratings via the PTSD Checklist-Military scale (Blanchard, Jones-Alexander, Buckley, & Forneris, 1996; Bliese et al., 2008), the active and control groups had similar scores" comment: The administered scale (PCL-M) is a validated PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "All subjective and objective laboratory assessments for the active group were conducted within 1 week before (Time 1) and 1 week after(Time2)the 7-day intervention" comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	comment: the assessors were the participants themselves (self-report instrument) who were probably aware of their group status
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: assessment could have been influenced because participants were not blind to their condition and might have answered according to their beliefs/expectations regarding the treatment efficacy or according to their beliefs about desired results (to please the investigators; demand effects). the risk of bias due to knowledge of the intervention is particularly high as the comparator is a passive control condition.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	the risk of bias due to knowledge of the intervention is high as participants were not blinded, a self-report measure was used, and the comparator was a no-treatment control condition.
	5.1 Were the data that produced this result analysed in accordance with a pre-specified	NI	

Bias in selection of the reported result	analysis plan that was finalized before unblinded outcome data were available for analysis?		
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	
	5.3 ... multiple eligible analyses of the data?	NI	
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure(s) of interest; for all time points of interest; results for the ITT sample reported; the risk of bias due to selection of results based on multiple eligible analyses might be lowered by the fact that the results assessed here are raw values (means, SDs) and data was available for almost all participants. All in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are some concerns.
Overall bias	**Risk of bias judgement**	**High**	computer-generated randomization list; insufficient information regarding allocation concealment; no baseline differences. the risk of bias due to knowledge of the intervention is high as participants were not blinded, a self-report measure was used, and the comparator was a no-treatment control condition. ; not enough information regarding pre-specified analysis plan to reliably assess the risk of bias due to selection of the reported result.

Unique ID	133	Study ID	1420101	Assessor	R
Ref or Label	Shalev 2012	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	PE	Comparator	CT	Source	Journal article(s) with results of the trial; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	PY	Quote: "The equipoise-stratified randomization is a method for randomly allocating participants to interventions intreatment studies that include more than 2 arms. It allows potential participants to decline treatment options that they do not desire and to be randomly assigned to the remaining arms. By making that choice, each participant assigns himself or herself to a "stratum,"which consists of all the options that he or she finds equally acceptable" "In our study, participants who agreed to start treatment (n=296)were informed about the 4 treatment options (PE, CT, treatment with SSRI vs placebo, and WL and subsequent delayed PE), could decline up to 2 treatment options (including the WL for delayed PE), and were randomly assigned to the remaining treatment options"
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "The study groups were similar with respect to age, traumatic events, and time to first treatment session (Table 1).There were more female participants in the CT group than in the other groups (P<.03), and there were higher PSS-SR scores in the SSRI group than in the other groups (P<.02)."
	Risk of bias judgement	**Some concerns**	equipoise-stratified randomization allowing participants to decline treatment options, which may bias the result (domain 2); no information on allocation concealment; no substantial baseline differences; all in all, there are concerns.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	Quote: "In our study, participants who agreed to start treatment (n=296)were informed about the 4 treatment options (PE, CT, treatment with SSRI vs placebo, and WL and subsequent delayed PE), could decline up to 2 treatment options (including the WL for delayed PE), and were randomly assigned to the remaining treatment options"
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	

			comment: therapists and participants were probably necessarily aware of the assigned condition as blinding is not feasible when psychological interventions are implemented especially not in view of the detailed elucidation regarding treatment conditions (quote)
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PY	Quote: "In our study, participants who agreed to start treatment (n=296)were informed about the 4 treatment options (PE, CT, treatment with SSRI vs placebo, and WL and subsequent delayed PE), could decline up to 2 treatment options (including the WL for delayed PE), and were randomly assigned to the remaining treatment options"
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	PY	comment: as the study is unblinded, the process of allowing participants to decline upto 50% of the treatment options and being randomized to a more acceptable treatment condition may have influenced the results (e.g. higher adherence rates)
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	Y	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	PY	Quote: "Because those who conducted the CA-2 and CA-3 were blinded to treatment attendance and adherence, the resulting comparisons include completers, partial completers, and noncompleters and thereby represent the total yield of participants randomly assigned to an intervention" "To account for missing observations and the groups' heterogeneities, we used a linear mixed model"
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Some concerns**	as the study is unblinded, the process of allowing participants to decline upto 50% of the treatment options and being randomized to a more acceptable treatment condition may have biased the results; ITT analysis using LMM.

Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "The study groups had similar retention rates between CA-1 and CA-2: 56 of 63 participants who received PE (88.9%), [...] 33 of 40 participants who received CT (82.5%), [...] and 79 WL participants (84.9%)."
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	comment: no analysis correcting for bias or sensitivity analysis
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: "The study groups had similar retention rates between CA-1 and CA-2"
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	comment: overall amount of missing posttest data =13.6%; equal proportions of dropout; no analysis conducted to identify potential differences between participants with and without missing posttest data; reasons for dropout are not reported; therefore, the risk of bias is high.
	Risk of bias judgement	**High**	missing posttest data 13.6%; equal proportions of dropout; no analysis conducted to identify potential differences between participants with and without missing posttest data; reasons for dropout not reported; therefore, the risk of bias is high.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The Clinicians-Administered PTSD Scale (CAPS) was used to confer a diagnosis of PTSD and a continuous measure of PTSD symptoms" comment: The administered scale (CAPS) is a validated, gold-standard PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points

			Quote: "The clinical assessments were made by clinical psychology interns. [...] They remained blind to treatment attendance and adherence."
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	comment: it is not clear whether "blind to treatment attendance and adherence" means that they were also blind to treatment allocation. Since participants were unblinded there was no blind assessment either way.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: The risk of bias due to lack of participant blinding is lowered by the fact that the comparator was also an active intervention (and, in addition, sharing common elements of psychotherapeutic treatment); however, it is not clear from the report whether the interviewer might have known the participants' allocation status. For subjective outcomes such as 'clinical impression of improvement', knowledge of the intervention received could be highly influential. Therefore, there are strong concerns.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN	
	Risk of bias judgement	**Some concerns**	risk of bias due to lack of participant blinding is lowered by the fact that the comparator was also an active intervention; there are concerns in view of the fact that interviewers might have been unbinded, too, which would increase the risk of bias.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: "Trial Registration: clinicaltrials.gov Identifier: NCT00146900" comment: The trial was prospectively registered (clinicaltrials.gov). examination of the history of changes and the comparison of study record versions indicates that authors adhered to their pre-specified intentions in all aspects that can be examined from the registry and that are relevant for the result of interest, except for one change: in the trial registration report it says that participants would be allowed to decline one treatment

			whereas in the paper it is reported that participants were allowed to decline upto two treatments.
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: pre-registered information is consistent with the final report concerning measures that were used; results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	
	Risk of bias judgement	**Some concerns**	The trial was prospectively registered at ClinicalTrials.gov. A comparison of this data with the reported information indicates that, all in all, researchers adhered to their pre-specified intentions. the trial registry offers no information regarding the statistical analysis plan, so it is difficult to assess whether the numerical result was selected on the basis of the results from multiple eligible analyses; based on the available information, all in all, it is not considered to be likely.
Overall bias	**Risk of bias judgement**	**High**	equipoise-stratified randomization procedure allowing participants to decline treatment options, which may bias the result (domain 2); no information on allocation concealment; no substantial baseline differences; all in all, there are concerns. as the study is unblinded, the process of allowing participants to decline upto 50% of the treatment options and being randomized to a more acceptable treatment condition may have biased the results; ITT analysis using LMM. missing posttest data 13.6%; equal proportions of dropout; no analysis conducted to identify potential differences between participants with and without missing posttest data; reasons for dropout not reported; therefore, there is a risk of bias. risk of bias due to lack of participant blinding is lowered by the fact that the comparator was also an active intervention; there are concerns in view of the fact that interviewers might have been

unbinded, too, which would increase the risk of bias. The trial was pre-registered at ClinicalTrials.gov; the information in the registry is consistent with the information reported in the paper in relevant aspects.

Unique ID	134	Study ID	1420102	Assessor	R
Ref or Label	Shalev 2012	**Aim**	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	PE	**Comparator**	WL	**Source**	Journal article(s) with results of the trial; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
Outcome	clinician-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	PY	Quote: "The equipoise-stratified randomization is a method for randomly allocating participants to interventions intreatment studies that include more than 2 arms. It allows potential participants to decline treatment options that they do not desire and to be randomly assigned to the remaining arms. By making that choice, each participant assigns himself or herself to a "stratum,"which consists of all the options that he or she finds
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	

			equally acceptable" "In our study, participants who agreed to start treatment (n=296)were informed about the 4 treatment options (PE, CT, treatment with SSRI vs placebo, and WL and subsequent delayed PE), could decline up to 2 treatment options (including the WL for delayed PE), and were randomly assigned to the remaining treatment options"
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "The study groups were similar with respect to age, traumatic events, and time to first treatment session (Table 1).There were more female participants in the CT group than in the other groups (P<.03), and there were higher PSS-SR scores in the SSRI group than in the other groups (P<.02)."
	Risk of bias judgement	**Some concerns**	equipoise-stratified randomization allowing participants to decline treatment options, which may bias the result (domain 2); no information on allocation concealment; no substantial baseline differences; all in all, there are concerns.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	Quote: "In our study, participants who agreed to start treatment (n=296)were informed about the 4 treatment options (PE, CT, treatment with SSRI vs placebo, and WL and subsequent delayed PE), could decline up to 2 treatment options (including the WL for delayed PE), and were randomly assigned to the remaining treatment options"
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as blinding is not feasible when psychological interventions are implemented especially not in view of the detailed elucidation regarding treatment conditions (quote)

	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PY	Quote: "In our study, participants who agreed to start treatment (n=296)were informed about the 4 treatment options (PE, CT, treatment with SSRI vs placebo, and WL and subsequent delayed PE), could decline up to 2 treatment options (including the WL for delayed PE), and were randomly assigned to the remaining treatment options"
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	PY	comment: as the study is unblinded, the process of allowing participants to decline upto 50% of the treatment options and being randomized to a more acceptable treatment condition may have influenced the results (e.g. higher adherence rates)
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	Y	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	PY	Quote: "Because those who conducted the CA-2 and CA-3 were blinded to treatment attendance and adherence, the resulting comparisons include completers, partial completers, and noncompleters and thereby represent the total yield of participants randomly assigned to an intervention" "To account for missing observations and the groups' heterogeneities, we used a linear mixed model"
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Some concerns**	as the study is unblinded, the process of allowing participants to decline upto 50% of the treatment options and being randomized to a more acceptable treatment condition may have biased the results; ITT analysis using LMM.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "The study groups had similar retention rates between CA-1 and CA-2: 56 of 63 participants who received PE (88.9%), [...] 33 of 40 participants who received CT (82.5%), [...] and 79 WL participants (84.9%)."
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	comment: no analysis correcting for bias or sensitivity analysis

	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: "The study groups had similar retention rates between CA-1 and CA-2"
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	comment: overall amount of missing posttest data =8.65%; equal proportions of dropout; no analysis conducted to identify potential differences between participants with and without missing posttest data; reasons for dropout are not reported; therefore, the risk of bias is high.
	Risk of bias judgement	**High**	missing posttest data 8.65%; equal proportions of dropout; no analysis conducted to identify potential differences between participants with and without missing posttest data; reasons for dropout not reported; therefore, the risk of bias is high.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The Clinicians-Administered PTSD Scale (CAPS) was used to confer a diagnosis of PTSD and a continuous measure of PTSD symptoms" comment: The administered scale (CAPS) is a validated, gold-standard PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "The clinical assessments were made by clinical psychology interns. [...] They remained blind to treatment attendance and adherence." comment: it is not clear whether "blind to treatment attendance and adherence" means that they were also blind to treatment

			allocation. Since participants were unblinded there was no blind assessment either way.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is rated higher because the comparator is a no-treatment condition.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	In addition, it is not clear from the report whether the interviewer was aware of the participants' allocation status. For subjective outcomes such as 'clinical impression of improvement', knowledge of the intervention received could be highly influential. Therefore, there are strong concerns.
	Risk of bias judgement	**High**	the risk of bias due to knowledge of the intervention is high as participants were not blinded and the comparator was a no-treatment control condition; also, unclear blinding of interviewers.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: "Trial Registration: clinicaltrials.gov Identifier: NCT00146900" comment: The trial was pre-registered (clinicaltrials.gov). examination of the history of changes and the comparison of study record versions indicates that authors adhered to their pre-specified intentions in all aspects that can be examined from the registry and that are relevant for the result of interest, except for one change: in the trial registration report it says that participants would be allowed to decline one treatment whereas in the paper it is reported that participants were allowed to decline upto two treatments.
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: pre-registered information is consistent with the final report concerning measures that were used; results were reported for all outcome measures of interest; for all time points of interest

	5.3 ... multiple eligible analyses of the data?	PN	
	Risk of bias judgement	**Some concerns**	The trial was prospectively registered at ClinicalTrials.gov. A comparison of this data with the reported information indicates that, all in all, researchers adhered to their pre-specified intentions. the trial registry offers no information regarding the statistical analysis plan, so it is difficult to assess whether the numerical result was selected on the basis of the results from multiple eligible analyses; based on the available information, all in all, it is not considered to be likely.
Overall bias	**Risk of bias judgement**	**High**	equipoise-stratified randomization procedure allowing participants to decline treatment options, which may bias the result (domain 2); no information on allocation concealment; no substantial baseline differences; all in all, there are concerns. as the study is unblinded, the process of allowing participants to decline upto 50% of the treatment options and being randomized to a more acceptable treatment condition may have biased the results; ITT analysis using LMM. missing posttest data 8.65%; equal proportions of dropout; no analysis conducted to identify potential differences between participants with and without missing posttest data; reasons for dropout not reported; therefore, there is a risk of bias. the risk of bias due to knowledge of the intervention is high as participants were not blinded and the comparator was a no-treatment control condition; also, unclear blinding of interviewers. The trial was pre-registered at ClinicalTrials.gov; the information in the registry is consistent with the information reported in the paper in relevant aspects.

Unique ID	135	Study ID	1420103	Assessor	R
Ref or Label	Shalev 2012	**Aim**	assignment to intervention (the 'intention-to-treat' effect)		

Experimental	CT	Comparator	WL	Source	Journal article(s) with results of the trial; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	PY	Quote: "The equipoise-stratified randomization is a method for randomly allocating participants to interventions intreatment studies that include more than 2 arms. It allows potential participants to decline treatment options that they do not desire and to be randomly assigned to the remaining arms. By making that choice, each participant assigns himself or herself to a "stratum,"which consists of all the options that he or she finds equally acceptable"
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	"In our study, participants who agreed to start treatment (n=296)were informed about the 4 treatment options (PE, CT, treatment with SSRI vs placebo, and WL and subsequent delayed PE), could decline up to 2 treatment options (including the WL for delayed PE), and were randomly assigned to the remaining treatment options"
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "The study groups were similar with respect to age, traumatic events, and time to first treatment session (Table 1).There were more female participants in the CT group than in the other groups (P<.03), and there were higher PSS-SR scores in the SSRI group than in the other groups (P<.02)."
	Risk of bias judgement	**Some concerns**	equipoise-stratified randomization allowing participants to decline treatment options, which may bias the result (domain 2); no information on allocation concealment; no substantial baseline differences; all in all, there are concerns.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	Quote: "In our study, participants who agreed to start treatment (n=296)were informed about the 4 treatment options (PE, CT, treatment with SSRI vs placebo, and WL and subsequent delayed PE), could decline up to 2 treatment options (including the WL for delayed PE), and were randomly assigned to the remaining
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	

			treatment options"
			comment: therapists and participants were probably necessarily aware of the assigned condition as blinding is not feasible when psychological interventions are implemented
			especially not in view of the detailed elucidation regarding treatment conditions (quote)
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PY	Quote: "In our study, participants who agreed to start treatment (n=296)were informed about the 4 treatment options (PE, CT, treatment with SSRI vs placebo, and WL and subsequent delayed PE), could decline up to 2 treatment options (including the WL for delayed PE), and were randomly assigned to the remaining treatment options"
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	PY	comment: as the study is unblinded, the process of allowing participants to decline upto 50% of the treatment options and being randomized to a more acceptable treatment condition may have influenced the results (e.g. higher adherence rates)
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	Y	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	PY	Quote: "Because those who conducted the CA-2 and CA-3 were blinded to treatment attendance and adherence, the resulting comparisons include completers, partial completers, and noncompleters and thereby represent the total yield of participants randomly assigned to an intervention"
			"To account for missing observations and the groups' heterogeneities, we used a linear mixed model"
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	

	Risk of bias judgement	**Some concerns**	as the study is unblinded, the process of allowing participants to decline upto 50% of the treatment options and being randomized to a more acceptable treatment condition may have biased the results; ITT analysis using LMM.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "The study groups had similar retention rates between CA-1 and CA-2: 56 of 63 participants who received PE (88.9%), [...] 33 of 40 participants who received CT (82.5%), [...] and 79 WL participants (84.9%)."
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	comment: no analysis correcting for bias or sensitivity analysis
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: "The study groups had similar retention rates between CA-1 and CA-2"
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	comment: overall amount of missing posttest data =15.79%; equal proportions of dropout; no analysis conducted to identify potential differences between participants with and without missing posttest data; reasons for dropout are not reported; therefore, the risk of bias is high.
	Risk of bias judgement	**High**	missing posttest data 15.79%; equal proportions of dropout; no analysis conducted to identify potential differences between participants with and without missing posttest data; reasons for dropout not reported; therefore, the risk of bias is high.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The Clinicians-Administered PTSD Scale (CAPS) was used to confer a diagnosis of PTSD and a continuous measure of PTSD symptoms" comment: The administered scale (CAPS) is a validated, gold-standard PTSD measure and likely to be sensitive to intervention effects

	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "The clinical assessments were made by clinical psychology interns. [...] They remained blind to treatment attendance and adherence." comment: it is not clear whether "blind to treatment attendance and adherence" means that they were also blind to treatment allocation. Since participants were unblinded there was no blind assessment either way.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is rated higher because the comparator is a no-treatment condition. In addition, it is not clear from the report whether the interviewer was aware of the participants' allocation status. For subjective outcomes such as 'clinical impression of improvement', knowledge of the intervention received could be highly influential. Therefore, there are strong concerns.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	the risk of bias due to knowledge of the intervention is high as participants were not blinded and the comparator was a no-treatment control condition; also, unclear blinding of interviewers.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: "Trial Registration: clinicaltrials.gov Identifier: NCT00146900"comment: The trial was pre-registered (clinicaltrials.gov). examination of the history of changes and the comparison of study record versions indicates that authors adhered to their pre-specified intentions in all aspects that can be examined from the registry and that are relevant for the result of interest, except for one change: in the trial registration report it says that participants would be allowed to decline one treatment whereas in the paper it is reported that participants were allowed to decline upto two treatments.

	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: pre-registered information is consistent with the final report concerning measures that were used; results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	
	Risk of bias judgement	**Some concerns**	The trial was prospectively registered at ClinicalTrials.gov. A comparison of this data with the reported information indicates that, all in all, researchers adhered to their pre-specified intentions. the trial registry offers no information regarding the statistical analysis plan, so it is difficult to assess whether the numerical result was selected on the basis of the results from multiple eligible analyses; based on the available information, all in all, it is not considered to be likely.
Overall bias	**Risk of bias judgement**	**High**	equipoise-stratified randomization procedure allowing participants to decline treatment options, which may bias the result (domain 2); no information on allocation concealment; no substantial baseline differences; all in all, there are concerns. as the study is unblinded, the process of allowing participants to decline upto 50% of the treatment options and being randomized to a more acceptable treatment condition may have biased the results; ITT analysis using LMM. missing posttest data 15.79%; equal proportions of dropout; no analysis conducted to identify potential differences between participants with and without missing posttest data; reasons for dropout not reported; therefore, there is a risk of bias. the risk of bias due to knowledge of the intervention is high as participants were not blinded and the comparator was a no-treatment control condition; also, unclear blinding of interviewers. The trial was pre-registered at ClinicalTrials.gov; the information in the registry is consistent with the information reported in the paper in relevant aspects.

Unique ID	141	Study ID	1480101	Assessor	R

Ref or Label	Suris 2013	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	CPT	Comparator	PCT	Source	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	N	Quote: "For the purpose of randomization, participants were assigned sequential PIN numbers as they entered the study. Blocks of random numbers were generated for each therapist, and were allocated to either CPT or PCT using a conditional statement. The random number sequence was maintained on an Excel spreadsheet, and as subjects' PINs were entered into the spreadsheet, the preassigned condition was revealed."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	N	comment: exclusion of eligible participants post-randomization, so the "ITT" sample assessed here is not the same sample that underwent randomization
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	NI	Quote: "There were no significant differences between the original sample of 129 and the final sample of 86 on demographic or baseline measures." "There were no significant differences between the CPT and PCT groups in self-reported PTSD severity scores at baseline (p = .85), [...]" comment: results of analyses of baseline characteristics are only

			reported for the modified sample of n=86 (original ITT sample n=129) after removing participants that had been treated by a therapist with low fidelity ratings. although it is stated that there were no differences between the samples, there are concerns
	Risk of bias judgement	**Some concerns**	insufficient information on randomization process with respect to allocation concealment and baseline differences in the original ITT sample.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as blinding is not feasible when psychological interventions are implemented
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: the reported deviations (which are likely to have affected the result) did not arise because of the trial context (e.g. failure of implementation of one therapist that led to the exclusion of 43 participants treated by this therapist)
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	

	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	N	Quote: "Due to treatment fidelity issues described in the Treatment Conditions subsection below, data were analyzed for N=86 participants. There were no significant differences between the original sample of 129 and the final sample of 86 on demographic or baseline measures." "Mixed-effects analysis of the intent-to-treat sample (N=86) [...]" comment: GLMM analysis to account for missing data; however, although the excerpt quoted indicates that analysis were conducted on data of the ITT sample this was not the case as eligible participants that had been randomized wer excluded post-treatment due to fidelity issues
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	PY	comment: considerable number of eligible participants (n=43; =33.33%) excluded from analysis
	Risk of bias judgement	**High**	analysis using GLMM; considerable number of eligible participants (n=43; =33.33%) excluded post-randomization due to fidelity issues.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "The combined treatment dropout rate was 28% (n=24), with rates of approximately 35% for CPT and approximately 18% for PCT."
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: "Although not statistically different from the 18% dropout rate of the PCT group in the current study, the 35% dropout rate in the CPT group was higher than in other randomized control trials of CPT" "Dropouts did not differ from treatment completers on baseline measures. Separate analyses of PTSD outcomes comparing treatment completers with dropouts were not performed as only 20.8% (n=5) of the treatment dropouts in the current study
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PY	

			participated in any of the four follow-up assessments."
			"In the current study, the vast majority of CPT dropouts occurred in the first half of treatment, specifically between sessions 3 and 6. Although we did not collect data on causes of dropout, we hypothesize that this is related to the fact that participants write their trauma narratives as homework for sessions 3 and 4."
			"CPT is also more demanding in terms of the mental focus and the homework required (as opposed to the traumafree journaling of PCT), which may have also contributed to the differential dropout between the treatment groups."
			"Dropout rates among the three therapists with acceptable fidelity did not significantly differ by therapist"
			"During the course of the study there were three adverse events in the CPT condition (one suicide attempt by overdose and two psychiatric hospitalizations)and two adverse events in the PCT condition (one suicide attempt by overdose and one psychiatric hospitalization). No events were deemed definitely studyrelated; however, one psychiatric hospitalization in the CPT condition was deemed possibly related."
			comment: difference in proportions of dropouts not statistically significant but potentially treatment-related as a temporal connection between dropout rates and homework content in CPT suggests; reasons for dropout are unknown; potential worsening of symptoms not assessed in dropouts; taken together, there is a significant risk of bias.
	Risk of bias judgement	**High**	difference in proportions of dropouts not statistically significant but missingness in the outcome is potentially treatment-related as the

			temporal connection between dropout rates and homework content in CPT suggests; reasons for dropout are unknown.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The CAPS (Blake et al., 1995) was used to assess current PTSD diagnosis and symptom severity." comment: The administered scale (CAPS) is a validated, gold-standard PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	NI	Quote: " Follow-up assessments including the CAPS, PCL, and QIDS [...]. Assessors were blind to treatment condition." comment: interviewers were blinded; nonetheless, no blind assessment since participants might have been aware of their allocation status; also, it was not systematically assessed whether blinding of interviewers was successful
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the outcome is not based solely on self-ratings and by the fact that the comparator was also an active intervention - which makes it less likely that participants' answers differed based on differential beliefs/expectations about the intervention effect
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN	
	Risk of bias judgement	**Some concerns**	the risk of bias due to knowledge of the intervention is lowered by the fact that the outcome is not based solely on self-ratings and by the fact that the comparator was also an active intervention

Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: "We hypothesized that both self-reported and clinician-assessed PTSD symptoms, as well as depressive symptoms, would show significantly greater reduction in the CPT group than in the PCT group. Only the hypothesis regarding self-reported PTSD symptoms was supported by the mixed-effects model analyses." comment: overall, authors offer detailed information on many potential sources of bias, report comprehensive justifications for changes in the analysis plan and include careful considerations regarding the risk of bias in their discussion; therefore, the risk of bias due to selective reporting is not considered high; however, with the lack of information in mind some concerns remain.
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest; non-significant results contrary to their hypotheses are also reported
	5.3 ... multiple eligible analyses of the data?	NI	
	Risk of bias judgement	**Some concerns**	authors offer detailed information on many aspects that pose potential sources of bias, report comprehensive justifications for changes in the analysis plan and include careful considerations regarding the risk of bias in their discussion; this suggests transparency; therefore, the risk of bias due to selective reporting is not considered high; however, with the lack of information in mind some concerns remain.

Overall bias	Risk of bias judgement	High	insufficient information on randomization process with respect to allocation concealment and baseline differences in the original ITT sample; slight concerns. analysis using GLMM; considerable number of eligible participants (n=43; =33.33%) excluded post-randomization due to fidelity issues. difference in proportions of dropouts not statistically significant but missingness in the outcome is potentially treatment-related as the temporal connection between dropout rates and homework content in CPT suggests; reasons for dropout are unknown. the risk of bias due to knowledge of the intervention is lowered by the fact that the outcome is not based solely on self-ratings and by the fact that the comparator was also an active psychotherapeutic intervention. although examination of the report indicates transparent reporting there is not enough information regarding the pre-specified analysis plan to reliably assess the risk of bias due to the selection of the reported result.

Unique ID	145	Study ID	1560101	Assessor	R
Ref or Label	van der Kolk 2014	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	MBI	Comparator	Psychoeducation	Source	Journal article(s) with results of the trial; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
	1.1 Was the allocation sequence random?	NI	

Bias arising from the randomizatio n process	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	Quote: "After an initial telephone screening, subjects were assessed, and, if eligible, randomly assigned to either trauma-informed yoga classes or women's health education classes" "Of the 83 participants, 7 (7%) withdrew consent prior to randomization and 12 (12%) withdrew consent prior to treatment; 64 (63%) were randomly assigned to treatment and formed the intention-to-treat (ITT) sample." comment: The statement in the latter quote is ambiguous with respect to dropouts; it is unclear whether (1) those 12 participants dropped out post-randomization (which would mean that the so called 'ITT sample' does not fulfill ITT principles) or (2) randomization was done a second time with the participants in the final sample (n=64) or (3) the randomization process described included only those 64 participants and the statement is somewhat misleading. This uncertainty raises some concerns.
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "Participants in the 2 treatment conditions did not differ significantly on any demographic variable (Table 1) or in any baseline measure of psychopathology, with the exception of significance in the employment demographic."
	Risk of bias judgement	**Some concerns**	insufficient information regarding random sequence generation and the sample; no information on allocation concealment; no substantial baseline differences on any demographic or clinical variables.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as blinding is not feasible in the context of this trial
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	

	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no deviations from the intended intervention reported that arose because of the trial context
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "We used hierarchical linear and nonlinear modeling with restricted maximum likelihood estimation to conduct multilevel regression analyses to examine change over time in outcomes as a function of treatment condition. This approach allowed us to analyze the intention-to-treat (ITT) sample without the use of missing data algorithms."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	open trial/no blinding; ITT analysis.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	PN	Quote: "Four people dropped out during the 10-week treatment phase, leaving 60 completers. There were no significant differences in dropout rates between the treatment groups, yoga (n = 1, 1.6%) and control (n = 3, 4.7%)." comment: amount of missing data in the yoga group =3.12; in the PsychEd group =9.37%
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: "There were no significant differences in dropout rates between the treatment groups, yoga (n = 1, 1.6%) and control (n = 3, 4.7%). There also were no significant differences between completers and dropouts on any baseline measure of psychopathology."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	

			comment: equal proportions of missing data between groups; no differences between participants with and without missing data on clinical measures at baseline; reasons for dropout not reported which leaves some uncertainty in the assessment; the overall amount of missing data is small
	Risk of bias judgement	**Some concerns**	equal proportions of missing data between groups; no differences between participants with and without missing data on clinical measures at baseline; reasons for dropout not reported which leaves some uncertainty in the assessment; the overall amount of missing data is small; taken together the risk of bias due to missing outcome data is not considered to be high although there are some concerns.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "Assessments were conducted at pretreatment, midtreatment (week 5), and posttreatment (week 10) and included the CAPS [...]" comment: The administered scale (CAPS) is a validated, gold-standard PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "All raters were blind to treatment condition." comment: : interviewers were blinded; nonetheless, no blind assessment since participants were probably aware of the assigned condition as blinding is not feasible in the context of this trial
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: assessment could have been influenced because participants were not blind to their condition and might have

	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN	answered according to their beliefs/expectations. The risk of bias due to lack of participant blinding, however, is lowered by the fact that the comparator was also an active intervention.
	Risk of bias judgement	**Some concerns**	interviewers blinded but participants unblinded; the risk of bias due to participants' knowledge of the intervention is lowered by the fact that the comparator was also an active intervention.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: "The study was registered on ClinicalTrials. gov (identifier: NCT00839813)." comment: The trial was pre-registered (clinicaltrials.gov). examination of the history of changes and the comparison of study record versions indicates that authors adhered to their pre-specified intentions in all aspects that are presented and that are relevant for the result of interest, except that the DTS as a secondary outcome measure of PTSD was added (at least it was not listed in the trial registry record)
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: see 5.1 (information from trial registry record)
	5.3 ... multiple eligible analyses of the data?	NI	
	Risk of bias judgement	**Some concerns**	The trial was prospectively registered at ClinicalTrials.gov; A comparison of this data with the reported information indicates that, all in all, researchers adhered to their pre-specified intentions; the uncertainty about the ITT sample (i.e. potential exclusion of eligible participants post-randomization) and subsequent ITT analyses raise slight concerns regarding data analysis.

Overall bias	Risk of bias judgement	Some concerns	insufficient information regarding random sequence generation; no information on allocation concealment; no substantial baseline differences on any demographic or clinical variables. amount of missing data =6.25%, equal proportions of dropout between groups; no differences between participants with and without missing data on clinical measures at baseline; reasons for dropout not reported which is why there are some concerns. interviewers blinded but participants unblinded; the risk of bias due to participants' knowledge of the intervention is lowered by the fact that the comparator was also an active intervention. The trial was pre-registered at ClinicalTrials.gov; A comparison of this data with the reported information indicates that, all in all, researchers adhered to their pre-specified intentions; the uncertainty about the ITT sample (i.e. potential exclusion of eligible participants post-randomization) and subsequent ITT analyses raise slight concerns regarding data analysis producing the reported result.

Unique ID	153	Study ID	1280101	Assessor	R
Ref or Label	Roth 2014	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	Relaxation	Comparator	WL	Source	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
	1.1 Was the allocation sequence random?	PY	

Bias arising from the randomizatio n process	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	Quote: "Eligible participants were assigned to either the PTSD hyperarousal group or no-PTSD group by the supervising psychologist and then randomly assigned to either IT or WL by the research assistant or study coordinator. A weighted randomization procedure was 35% more likely to assign new participants to the IT than WL condition because of differing withdrawal rates."
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "No baseline demographic and clinical differences were found in the IT and WL PTSD groups"
	Risk of bias judgement	**Some concerns**	no information on allocation concealment; no baseline differences.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	Y	Quote: "We report here a nonblinded randomized clinical trial of CART in veterans with PTSD hyperarousal symptoms."
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	"Diagnostic raters were not blind to group membership, because participants themselves typically revealed their treatment status"
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "Primary treatment outcomes from baseline to first follow-up in the intent-to-treat PTSD sample were analyzed by mixed modeling using restricted maximum likelihood estimation."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no deviations from the intended intervention that arose because of the experimental context; ITT analysis using mixed modeling.
	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "Fifty-one participants across the PTSD groups (63.8%) completed Assessment Time Point 2, which was used for the primary outcome analyses."

Bias due to missing outcome data	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: " The most common reasons cited for study withdrawal were inconvenience and study burden." "Did not complete breathing training (n=22) - intervention too much work/time (n=16) - moved out of area (n=3) - lost to follow-up (n=2) - withdrew when study did not provide letter attesting PTSD diagnosis (n=1)" "Did not complete waiting period (n=6) - moved out of area (n=3) - lost to follow-up (n=3)"
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PY	"Most of the withdrawals occurred before the first treatment session. At that point, participants were aware of the tasks and time required for the study but not the treatment rationale. We did not assess treatment credibility until after the first treatment session, so we can only speculate that some participants may have discontinued beforehand, believing the treatment approach was implausible." comment: amount of missing data: overall = 36.25%, in REL =48.93%, in WL =18.18%, difference in proportions between treatment groups [no significance test]; reasons for missing posttest data may be related to the assigned treatment in n=19 cases in the Relaxation group and in n=3 cases in the WL group; reasons for dropout differ between groups; it is unknown whether participants with and without missing posttest data differed on any study variables; all in all, the missingness mechanism can be assumed to be non-ignorable.
	Risk of bias judgement	**High**	amount of missing data: overall = 36.25%, in REL =48.93%, in WL =18.18%, difference in proportions of missing data across groups;

			reasons for dropout differ between groups; it is likely that missingness in the outcome depended on its true value.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "CAPS-IV total and CAPS-IV hyperarousal subscale were highly reliable in our sample, with Cronbach alphas of .97 and .92, respectively. Diagnostic interviewers attained 100% agreement on all SCID diagnoses, and CAPS-IV interrater reliability coefficients were greater than .90 on training cases prior to interviewing new participants." comment: The administered scale (CAPS) is a validated, gold-standard PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	Quote: " Diagnostic raters were not blind to group membership, because participants themselves typically revealed their treatment status during the Clinician-Administered PTSD Scale for DSM–IV (CAPS-IV; Blake et al., 1995) interview."
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is high because the comparator is a no-treatment condition.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	In addition, the interviewer might have known the participants' treatment condition. For subjective outcomes such as 'clinical impression of improvement', knowledge of the intervention received could be highly influential.

	Risk of bias judgement	**High**	risk of bias due to knowledge of the intervention: neither participants nor interviewers were blinded; The risk is particularly high for two reasons: (1) the comparator was a no-treatment waitlist condition increasing the probability that differential beliefs/expectations concerning the efficacy of both treatment conditions (or regarding desired study results) affected their answers and (2) for subjective outcomes such as 'clinical impression of improvement', knowledge of the intervention received could be highly influential.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: " This study was approved by the Stanford University Institutional Review Board and registered as Clinical Trial No. NCT00855816." comment: no statistical analysis plan (SAP) provided; but the available information on researchers' pre-specified intentions indicates that, overall, investigators adhered to their pre-specified intentions
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest; administered scales (as reported in the publication) are in line with those that had been specified in the trial registry record
	5.3 ... multiple eligible analyses of the data?	PN	comment: see 5.1
	Risk of bias judgement	**Low**	no statistical analysis plan (SAP) provided; but the available information on researchers' pre-specified intentions indicates that, overall, investigators adhered to their pre-specified intentions

Overall bias	Risk of bias judgement	High	no information on allocation concealment. amount of missing data: overall = 36.25%, in REL =48.93%, in WL =18.18%, difference in proportions of missing data across groups; reasons for dropout differ between groups; it is likely that missingness in the outcome depended on its true value. risk of bias due to knowledge of the intervention: neither participants nor interviewers were blinded; The risk is particularly high for two reasons: (1) the comparator was a no-treatment waitlist condition increasing the probability that differential beliefs/expectations concerning the efficacy of both treatment conditions (or regarding desired study results) affected their answers and (2) for subjective outcomes such as 'clinical impression of improvement', knowledge of the intervention received could be highly influential. no statistical analysis plan (SAP) provided; but the available information on researchers' pre-specified intentions indicates that, overall, investigators adhered to their pre-specified intentions.

Unique ID	156	Study ID	1690101	Assessor	R
Ref or Label	Goldstein 2017	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	MBI	Comparator	WL	Source	Journal article(s) with results of the trial; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
Outcome	clinician-rated PTSD	Results	SMD (post-post)	Weight	1

Domain	Signalling question	Response	Comments
	1.1 Was the allocation sequence random?	Y	

Bias arising from the randomization process	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	Quote: " Participants were randomized to either IE or WL, with randomization determined by blocked randomization lists from four strata defined by gender and age (18–50 or 51–69 years)."
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	Quote: "There were no significant differences between groups on any observed baseline variables." "Based on LSI scores, baseline exercise level did not differ between participants randomized to IE (M = 23.90, SD = 19.97) or WL (M = 26.00, SD = 26.32), t(42) = .29, p = .77." comment: no significance test of baseline differences between groups on CAPS scores (IE =64.25 (20.54), WL =58.5 (14.19))
	Risk of bias judgement	**Some concerns**	blocked randomization used; no information on allocation concealment; no substantial baseline differences.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as blinding is not feasible
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "These are presented as intent-to-treat analyses based on all available data. Mixed effects models were used to evaluate change in primary outcome measures using all available data, [...]"
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to	NA	

	analyse participants in the group to which they were randomized?		
	Risk of bias judgement	**Low**	no blinding, no deviations that arose because of the trial context; ITT analysis conducted.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: " Twenty-one veterans were randomly assigned to IE and 26 to WL." "Treatment completers in the IE group (n = 16) attended 28 in-person sessions [...]" comment: see Figure 1 flow chart: in IE group, 16 out of 21 completed post-test; in WL group, 22 out of 26 completed post-test
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis reported
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: "There was no significant difference in attrition between groups, with 5 participants (24%) discontinuing IE and 4 (15%) dropping from the waitlist condition, $\chi2(1) = .53$, $p = .47$. Most attrition occurred immediately or shortly after randomization; [...]"
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	"Baseline levels of the outcome variables appeared to be moderately related to number of sessions ultimately attended, with higher baseline PTSD symptom severity associated with greater number of sessions attended ($r = .26$, $p = .35$), and lower baseline physical quality of life ($r = −.41$, $p = .13$) and psychological quality of life ($r = −.53$, $p = .04$) also associated with greater number of sessions attended." comment: equal proportions of dropout between both groups; total amount of missing data: 19%; reasons for dropout are not reported, except for one WL participant who was "dissatisfied with

			randomization result"; it is unknown whether participants with and without post-test data differed on any relevant variable
	Risk of bias judgement	High	equal proportions of dropout: 24% missing data in IE, 15% in WL; no evidence that result was not biased; reasons for dropout unknown, as well as whether participants with and without missing post-test data differed on any relevant variables.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "Current and lifetime PTSD symptoms were assessed via CAPS interview based on the DSM-IV (American Psychiatric Association, 2000)" comment: The administered scale (CAPS) is a validated, gold-standard PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "The monitor only waitlist control condition lasted for 12 weeks, with participants completing study interviews and questionnaires at the same intervals as those in the IE condition." comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "Clinical interviewers were blind to treatment condition." comment: although the assessors were blind to the participants' condition, there was no blind assessment as participants were not blinded
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: interviewers were blinded; nonetheless, no blind assessment since participants might have been aware of their allocation status; (also, it was not systematically assessed whether blinding of interviewers was successful).
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	

			Because the comparator is a passive control condition the risk is higher that participants might have answered questions according to their beliefs/expectations regarding their assigned condition. This risk cannot be fully eliminated by assessment by a blinded clinician.
	Risk of bias judgement	**High**	the risk of bias due to knowledge of the intervention is high as participants were not blinded and the comparator was a no-treatment control condition.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: "This pilot randomized controlled trial for military veterans with PTSD was registered in a public registry (ClinicalTrials.gov identifier NCT01674244)." comment: The trial was pre-registered (clinicaltrials.gov) before recruitment. examination of the history of changes and the comparison of study record versions indicates that authors adhered to their pre-specified intentions in all aspects that are presented and that are relevant for the result of interest
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest; reported measures are in line with pre-specified intentions (registry record)
	5.3 ... multiple eligible analyses of the data?	PN	comment: results are reported for all participants randomized (ITT sample); no statistical analysis plan available, but the risk is considered relatively low since means, SDs and Ns are reported;
	Risk of bias judgement	**Low**	The trial was pre-registered at ClinicalTrials.gov. ; results were reported for the outcome measure(s) of interest; for all time points of interest; results for the ITT sample reported; comparison of data in registry record with information in the final report indicates that, all in all, researchers adhered to their pre-specified intentions. The risk of bias is therefore low.

Overall bias	Risk of bias judgement			High	blocked randomization used; no information on allocation concealment; no substantial baseline differences. no blinding, no deviations that arose because of the trial context; ITT analysis conducted. equal proportions of dropout: 24% missing data in IE, 15% in WL; no evidence that result was not biased; reasons for dropout unknown, as well as whether participants with and without missing post-test data differed on any relevant variables. the risk of bias due to knowledge of the intervention is high as participants were not blinded and the comparator was a no-treatment control condition. The trial was pre-registered at ClinicalTrials.gov. ; comparison of data in registry record with information in the final report indicates that, all in all, researchers adhered to their pre-specified intentions.

Unique ID	157	Study ID	1290101	Assessor	R
Ref or Label	Rothbaum 2012	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	PE	Comparator	WL	Source	Journal article(s) with results of the trial; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
Outcome	self-rated PTSD	Results	SMD	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "Envelopes containing computer-generated patient random assignments (either to immediate intervention or assessment only) were given to the patient and their nurse after the initial evaluation to ensure that assessors remained blind. The on-call therapist immediately provided the intervention to those assigned to this condition."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	

	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	Quote: "A logistic regression demonstrated no significant differences in baseline demographics between conditions. Univariate analyses of variance suggested that assessment and intervention conditions differed on ISRC numbing [F(1,135) = 4.39, p< .05] and ISRC reexperiencing [F(1,135) = 6.70, p< .01]."
	Risk of bias judgement	**Low**	computer-generated random allocation sequence; envelopes (sealed?!) used to ensure allocation concealment which indicates that probably the allocation sequence was concealed until participants were assigned; no substantial baseline differences that are unlikely to occur by chance.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	Y	Quote: "Envelopes containing computer-generated patient random assignments [...] were given to the patient and their nurse [...]"
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	comment: therapists and participants were probably necessarily aware of the assigned condition as blinding is not feasible when psychological interventions are implemented especially not in view of the obvious differences between an active and a passive treatment condition
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "Linear mixed-effect models were used to obtain predicted mean values for outcomes at each assessment point (weeks 4 and 12)."

			"Missing values for week 4 and week 12 data were handled with multiple imputation."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	no blinding; no deviations that arose because of the trial context; ITT analysis.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: " Of the 137 participants who were enrolled in the study, 102 (74%) completed 4-week follow-up [...]"
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	Quote: "Missing values for week 4 and week 12 data were handled with multiple imputation. The NORM (51) software package was used to generate 100 complete datasets in which demographic variables, pretreatment self-report measures, treatment condition, and trauma type were used as auxiliary variables."
			comment: no analysis correcting for bias or sensitivity analysis reported
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: Of the 137 participants who were enrolled in the study, 102 (74%) completed 4-week follow-up and 91 (66%) completed 12-week follow-up. No significant group differences in dropout rates were detected, X^2 =1.92, p =.17. No patients reported a desire to withdraw from the study as a result of their participation, and no study-related adverse effects were reported."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	comment: no information on group-level dropout rates, so it is unclear whether proportions of dropout differed; the amount of missing data is substantial; reasons for dropout are not reported; it

			is unknown whether participants with and without post-test data differed on any relevant variable.
	Risk of bias judgement	**High**	unknown whether proportions of missing post-test data differed between groups; total amount of missing data is substantial; reasons for dropout are not reported; unclear whether participants with and without post-test data differed on any relevant variables.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The PDS has high internal consistency. Test-retest reliability was good, from .74 to .85. High diagnostic agreement (82%) with the Structured Clinical Interview for DSM Disorders was noted (42)."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	comment: the PDS is a self-report questionnaire; hence the 'assessors' (the participants themselves) were probably aware of their intervention
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: assessment could have been influenced because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect;
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	the risk of bias due to knowledge of the intervention is high, as the comparator is a passive control condition (WL).
	Risk of bias judgement	**High**	self-report measure used; the risk of bias due to knowledge of the intervention is high as participants were not blinded and the comparator was a no-treatment control condition.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: " The study was registered at clinicaltrials. gov, NCT00895518." comment: The study started in 2008, however, the first registry entry is from 2009. Although registration was not done prior to enrollment, it was done before analyses, and information posted in the registry record in 2009 (and in later versions) is in line with

			data reported in the publication. Eligibility criteria were changed after study start - instead of including only females, all sexes were included.
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	comment: results were reported for all outcome measures of interest; for all time points of interest; reported measures are in line with intentions specified (regarding choice of measures) in the registry record (before analysis)
	5.3 ... multiple eligible analyses of the data?	PN	comment: results are reported for all participants randomized (ITT sample); no statistical analysis plan available, but the risk is considered relatively low since means, SDs and Ns are reported in addition to relative effect sizes
	Risk of bias judgement	**Low**	trial was registered at ClinicalTrials.gov.; results were reported for the outcome measure(s) of interest; for all time points of interest; results for the ITT sample reported; comparison of data in registry record with information in the final report indicates that, all in all, researchers adhered to their pre-specified intentions. The risk of bias is considered low
Overall bias	**Risk of bias judgement**	**High**	Randomization: envelopes (not reported whether they were sealed?) used to ensure allocation concealment which indicates that probably the allocation sequence was concealed until participants were assigned. Missing data: unknown whether proportions of missing post-test data differed between groups; total amount of missing data is substantial; reasons for dropout are not reported; unclear whether participants with and without post-test data differed on any relevant variables. Measurement: self-report measure used; the risk of bias due to

				knowledge of the intervention is high as participants were not blinded and the comparator was a no-treatment control condition.	

Unique ID	23	Study ID	210101	Assessor	R
Ref or Label	de Bont 2013	**Aim**	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	PE	**Comparator**	EMDR	**Source**	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	**Results**	SMD (post-post)	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	NI	Quote: "The enrollment phase (Fig. 1) was completed with the random assignment of the 10 patients to either PE or EMDR treatment."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	comment: no information regarding allocation concealment and no information on random sequence generation. Merely reporting that patients were randomly assigned without explicitly reporting that

			generation of the allocation sequence included a random element is not considered sufficient
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	NI	comment: baseline characteristics are reported for all 10 participants individually but without reporting group assignment, so it is not possible to detect potential differences between groups on any demographic or clinical measures. Results of clinical outcome measures are also only reported for all participants pooled together, not seperately for each treatment group
	Risk of bias judgement	**Some concerns**	no information regarding allocation concealment and no information on random sequence generation; baseline characteristics are reported for all 10 participants individually but without information on group status.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	NI	comment: therapists were probably nessessarily aware and participants might have been aware of their group status as blinding is not feasible. The fact that both groups received an active treatment (and interventions were similar in many aspects) reduces the probability that participants guessed their group status. However, it is not reported whether participants were informed about their assignment or not.
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "Second, using Wilcoxon pairwise tests, the CAPS PTSD total scores of the intention-to-treat (ITT) group at posttreatment (T2) and follow-up (T3) were compared with those at baseline (T1)"
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	
	Risk of bias judgement	**Low**	

Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "Drop out (n=2) PE (n=1) EMDR (n=1)" "Post-treatment measurements (n=8 -> 4 PE, 4 EMDR)" comment: amount of missing data at post-treatment T2 = 20%
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis reported
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NI	comment: equal proportions of dropout; not reported whether participants with and without posttest data differed on any demographic or clinical variables; reasons for dropout are not reported
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	
	Risk of bias judgement	**High**	missing posttest data =20%; equal proportions of dropout; not reported whether participants with and without posttest data differed on any demographic or clinical variables; reasons for dropout are not reported.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The CAPS (Blake et al., 1995; Dutch version: Hovens et al., 2005) was used to check the diagnostic criteria for PTSD and to assess the severity of PTSD symptoms. The CAPS rates the frequency and intensity of the DSM-IV-TR criteria; its total score ranges from 0 to 136. The reliability, validity, and sensitivity of the CAPS are good (Weathers, Keane, & Davidson, 2001; Dutch version: reliability alpha = .93 to .98; Hovens et al., 1994)"
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	
	4.3 Were outcome assessors aware of the intervention received by study participants?	NI	comment: therapists were probably nessessarily aware and participants might have been aware of their group status as blinding is not feasible. The fact that both groups received an active treatment (and interventions were similar in many aspects) reduces the probability that participants guessed their group status. However, it is not reported whether participants were informed about their assignment or not. No information on blindness of interviewers

	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention. However, unclear blinding of interviewers: If neither CAPS interviewers nor participants were blinded then risk of bias will be very high. For subjective outcomes such as 'clinical impression of improvement', knowledge of the intervention received could be highly influential.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NI	
	Risk of bias judgement	**High**	the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention; However, unclear blinding of interviewers; All in all, risk rated high due to lack of information.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	comment: results are reported for all outcome measures of interest at all time points of interest; however, numerical results are only reported for all participants pooled together which raises concerns; group-level numerical results not reported
	5.3 ... multiple eligible analyses of the data?	NI	comment: estimated marginal means and standard errors are reported (but not on group-level); no information regarding likeliness of selection of numerical results on the basis of multiple eligible analyses
	Risk of bias judgement	**Some concerns**	no information on statistical analysis plan; numerical results are only reported for all participants pooled together; group-level numerical results are not reported; all in all, difficult to assess RoB in this domain due to lack of information.

Overall bias	Risk of bias judgement			High	strong concerns regarding randomization process: no information on random sequence generation, allocation sequence, or baseline differences. missing posttest data =20%; equal proportions of dropout; not reported whether participants with and without posttest data differed on any demographic or clinical variables; reasons for dropout are not reported. the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention; However, unclear blinding of interviewers; All in all, risk rated high due to lack of information. no information on statistical analysis plan; numerical results are only reported for all participants pooled together which raises concerns; group-level numerical results are not reported; all in all, difficult to assess RoB in this domain due to lack of information.

Unique ID	158	Study ID	1500101	Assessor	R
Ref or Label	Taylor 2003	Aim	assignment to intervention (the 'intention-to-treat' effect)		
Experimental	EMDR	Comparator	REL	Source	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	Results	SMD (post-post)	Weight	1

Domain	Signalling question		Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?		NI	Quote: "Participants meeting study criteria were randomized to
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?		NI	eight 90-min individual sessions of either exposure therapy, EMDR, or relaxation training."

			comment: no details regarding randomization process, no explicit mentioning of use of a random component; no information on allocation concealment
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	NI	comment: no baseline characteristics on group-level reported
	Risk of bias judgement	**Some concerns**	insufficient information regarding random sequence generation; no information on allocation concealment; no information regarding potential baseline differences on demographic or clinical variables.
Bias due to deviations from intended interventions	2.1.Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably aware of the assigned condition as blinding is not feasible when psychological interventions are implemented
	2.2.Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	PN	comment: no effects of recruitment, engagement in activities on trial participants or personnel undermining the implementation of the trial protocol reported that might lead to bias
	2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?	NA	
	2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?	NA	
	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Y	Quote: "Intent-to-treat analyses for the four PTSD symptom dimensions were based on all 60 participants, using the last available treatment outcome assessment."
	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	NA	

	Risk of bias judgement	Low	no deviations that arose because of the trial context; ITT analysis used
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "The number of trial entrants and number of treatment completers were as follows: EMDR 19, 15; exposure therapy 22, 15; relaxation training 19, 15. The proportion of dropouts did not differ across treatments: $X^2(2, N = 60) = 0.86$, $p > .1$, $n^2 = .01$)."
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	comment: no analysis correcting for bias or sensitivity analysis reported
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	Quote: "Dropouts and completers did not differ on demographics, trauma
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	type, PTSD duration, or pretreatment scores on the primary or secondary outcome measures (ps >.05)."
			comment: equal proportions of dropout between both groups; the overall amount of missing data is substantial (=21%); reasons for dropout are not reported which raises some concerns; no differences between participants with and without posttest data on relevant variables
	Risk of bias judgement	Some concerns	no blinding; equal proportions of dropout; overall amount of missing data =21%; reasons for dropout not reported which raises some concerns; no differences between participants with and without posttest data.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "Severity of PTSD symptoms over the past week was assessed by the Clinician Administered PTSD subscales (CAPS; Blake et al., 1997)."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points

	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "All interviews were conducted by clinic staff, who were blind to the participants' treatment assignment." comment: interviewers were blinded; nonetheless, no blind assessment since participants might have been aware of their allocation status and might have answered interview questions according to their beliefs/expectations regarding the treatment effect
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	comment: assessment could have been influenced because participants were not blind to their condition and might have answered according to their beliefs/expectations regarding the treatment efficacy or according to their beliefs about desired results (to please the investigators; demand effects). This risk, however, is lowered by the fact that the comparator was also an active intervention.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN	
	Risk of bias judgement	**Some concerns**	blinded interviewers but unblinded participants; the risk of bias due to participants' knowledge of the intervention is lowered by the fact that the comparator was also an active intervention but some concerns remain.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	comment: insufficient information
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	comment: no information available
	5.3 ... multiple eligible analyses of the data?	PN	comment: results are reported for all participants randomized (ITT sample); no statistical analysis plan available, but the risk is considered relatively low since means, SDs and Ns are reported
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure of interest; for all time points of interest; All in all, it is difficult to assess the risk of bias in this domain due to lack of information; there are some concerns.

| Overall bias | Risk of bias judgement | Some concerns | insufficient information regarding random sequence generation; no information on allocation concealment; no information regarding potential baseline differences on demographic or clinical variables. no blinding; equal proportions of dropout; overall amount of missing data =21%; reasons for dropout not reported which raises some concerns; no differences between participants with and without posttest data. blinded interviewers but unblinded participants; the risk of bias due to participants' knowledge of the intervention is lowered by the fact that the comparator was also an active intervention but some concerns remain. not enough information regarding pre-specified analysis plan to reliably assess the risk of bias due to selection of the reported result. |

Summary of evaluations of the risk of bias of effect sizes based on analysis of study completers

Unique ID	1	Study ID	10101	Assessor	R
Ref or Label	Acarturk 2015	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention ...**	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	EMDR	**Comparator**	WL	**Source**	Journal article(s) with results of the trial
Outcome	self-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "For the allocation of participants to different treatment groups, a computer-generated random number list was used. Participants were randomly assigned on a 1:1 basis to the EMDRor wait-list group."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PY	Quote: "The sociodemographic data did not differ significantly between two groups (Table 1)." "At pretreatment, IES-R scoreswere significantly higher in the EMDR group than in the control group (EMDR group: M=64.80, SD=12.08 vs. wait-list group: M=56.93, SD=7.15), t (27)=2.115, p=0.044, d=0.76, 95% CI (0.01, 1.52)."
	Risk of bias judgement	**High**	computer-generated random sequence; no information on allocation concealment; preteatment PTSD scores sign. different
	2.1 Were participants aware of their assigned intervention during the trial?	Y	

	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	Quote: "The participants and the therapists were aware of the allocated arm, but the outcome assessors were kept blind to the allocation."
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	comment: no non-protocol interventions reported
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	Quote: "None of the participants gave permission for the video or audiotaping of the sessions. The reason reported for refusal was fear of the Syrian government. [...] The supervisor checked during live and normal one-on-one and group supervision sessions whether the therapists were complying with the 8 Phase EMDR Standard Protocol (Shapiro, 2001). Treatment fidelity was supported by the supervisor, who attended at least one session of each therapist."
Bias due to deviations from intended interventions	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PY	Quote: "In this study, a maximum of seven sessions (90 min per session) of EMDR were conducted." "The mean number of EMDR sessions was 4.13 (SD= 1.73, range 2-7)." comment: number of sessions attended differed between participants, the quote above indicates that this was in line with the intervention regimen, since a "maximum of 7 sessions" was intended to be delivered. However, in the trial registry record it is clearly stated that 7 sessions were intended to be delivered to participants. In view of the fact that there were no dropouts and participants received more than half of the sessions (59%) on average it is uncleat whether outcomes were significantly affected. nonetheless, there are concerns. (without adjusting, see 2.6) Non-attendance can potentially lead to bias towards the null
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	PN	comment: naive 'per protocol' analysis used; no inverse probability weighting to adjust for censoring

			of participants who cease adherence to their assigned intervention
	Risk of bias judgement	**Some concerns**	neither participants nor therapists blinded; no substantial deviations from intended interventions except that participants attended a mean of 4.13 sessions out of 7 sessions.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	Quote: "There was no dropout from treatment in the EMDR group, or during completion of the assessments (also WL)."
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	no missing data
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	PN	Quote: "The validity of IES-R has been tested in different populations (Panahi et al., 2011)." " Administration of the scale in a sample of native Arabic speakers (Zaghrout, 2013) yielded a Cronbach's alpha of a=0.93. The test-retest reliability calculated by administering the scale to the same sample on two occasions, 2 weeks apart, yielded a Pearson correlation coefficient of r=0.88 (Zaghrout, 2013)."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: same assessment times and procedures used for all participants
	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	Quote: "All questionnaires were self-report instruments, but for those who needed help, a research assistant who was blind to the treatment conditions administered the scales verbally."

			comment: the 'assessors' were -in this case- the participants themselves (who were aware of their intervention).
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is rated high since the comparator is a no-treatment condition
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NI	
	Risk of bias judgement	**High**	the risk of bias due to participants' knowledge of the intervention is rated high because the comparator is a no-treatment condition (here: WL)
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: "The study was registered to Clinical Trials (ClinicalTrials.gov Identifier: NCT01847742). [...] The consort checklist is available as supporting information(see Checklist_S1)." comment: examination of the History of Changes indicates that investigators adhered to their pre-specified intentions, except for the PTSD measure (see 5.2)
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PY	Quote from clinicaltrials.gov trial registry record: "Primary Outcome Measures: 1. score on Harvard Trauma Questionnaire (at the main study) [Time Frame: before and after the treatment,an expected average of 7 weeks of EMDR treatment] The change in HTQ score will be assessed after the EMDR treatment has finished after 7 weeks(estimated)."

			comment: both the HTQ and the IES are listed as primary outcome measures in the registry record, however, in the published report only the IES is mentioned and results are only reported for the IES. This raises concerns.
	5.3 ... multiple eligible analyses of the data?	PN	comment: raw values (means, SDs) (=unadjusted effect size estimates) are reported for all completers. (Baseline adjusted effect estimated (ANCOVA) were also reported)
	Risk of bias judgement	**High**	Trial registration on clinicaltrials.gov; both the HTQ and the IES are listed as primary outcome measures in the registry record, however, in the published report only the IES is mentioned and results are only reported for the IES.
Overall bias	**Risk of bias judgement**	**High**	no information on allocation concealment, difference in pretreatment scores. neither participants nor therapists blinded; no substatial deviations from intended interventions. no missing data. high risk of bias due to participants' knowledge of the intervention. Trial registration on clinicaltrials.gov; both the HTQ and the IES are listed as primary outcome measures in the registry record, however, in the published report only the IES is mentioned and results are only reported for the IES.

Unique ID	4	**Study ID**	30101	**Assessor**	R
Ref or Label	Adenauer 2011	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to**	occurance of non-protocol interventions; failures in implementing the intervention that could

				intervention...	have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	NET	**Comparator**	WL	**Source**	Journal article(s) with results of the trial; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
Outcome	clinician-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "Participants that fulfilled the inclusion criteria were randomized into the two groups using a computer-generated list of random numbers."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PY	Quote: "Baseline characteristics of the groups were compared to examine the effects of randomization using the MannWhitney U test for continuous variables and the ChiSquare test for dichotomous variables." comment: results of this analysis reported in table 3. no significance test is reported but given the large difference in pretreatment PTSD scores it is likely that the difference is statistically significant. In addition, there was a sign. difference in depression scores
	Risk of bias judgement	**High**	unclear allocation concealment; substantial differences between groups on clinical measures
Bias due to deviations from	2.1 Were participants aware of their assigned intervention during the trial?	Y	comment: therapists and participants were probably necessarily aware of the assigned condition
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	

intended interventions	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	Quote: "No major deviations from treatment protocol were detected. [...] Changes in medication status from pre- to posttest were little: At posttest, only two treated patients still took antidepressants (compared to four at pretest) and none still took anxiolytics (compared to one at pretest). At posttest, four WLC-patients were medicated with antidepressants (compared to two at pretest) and one was medicated with neuroleptics (compared to none at pretest). [...] However, as there were no differences between treatment groups with respect to intake of medication, [...]"
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	N	Quote: "No major deviations from treatment protocol were detected."
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	Quote: "No major deviations from treatment protocol were detected."
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	therapists and participants were probably unblinded; no major deviations
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	comment: 31 of 34 completed treatment and were assessed at t2. one dropout in NET, and two in WL (all dropped out unvoluntarily, due to deportation)
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PN	comment: 31 of 34 completed treatment and were assessed at t2. one dropout in NET, and two in WL (all dropped out unvoluntarily, due to deportation)
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	31 of 34 completed treatment and were assessed at t2. one dropout in NET, and two in WL (all dropped out unvoluntarily, due to deportation)
	4.1 Was the method of measuring the outcome inappropriate?	N	comment: measured with CAPS

	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	
Bias in measurement of the outcome	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "Posttests included the same instruments as used in the pretest and were carried out by interviewers who were blind to treatment condition." comment: comment: although the assessors were blind to the participants' condition, there was no blind assessment: Accoring to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: In the report there is no information indicating that the assessor was an independent researcher not involved in the study, which would otherwise lower the risk of bias.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NI	
	Risk of bias judgement	**High**	comment: the risk of bias due to knowledge of the intervention is rated higher if the comparator is no treatment (here: WL) than when the comparator is another active intervention. In addition, for subjective outcomes such as 'clinical impression of improvement', knowledge of the intervention received could be highly influential.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: "Registration of the clinical trial: Number: NCT00563888" comment: History of changes was examined. information consistent with information in the report
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest

	5.3 ... multiple eligible analyses of the data?			PN	comment: both ITT and completers data is available; generally the risk of bias due to mutiple eligible analyses of the data is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement			**Low**	Trial was registered. History of changes was examined. information consistent with information in the report
Overall bias	**Risk of bias judgement**			**High**	unclear allocation concealment; substantial baseline differences between groups on clinical measures. therapists and participants were probably unblinded; no major deviations. risk of bias due to knowledge of the intervention because participants were not blind to their condition. Trial was registered. History of changes was examined. information consistent with information in the report

Unique ID	8	**Study ID**	70102	**Assessor**	R
Ref or Label	Bormann 2008	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention ...**	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	MBI	**Comparator**	WL	**Source**	Journal article(s) with results of the trial; Personal communication with trialist
Outcome	clinician-rated PTSD	**Results**	SMD(between)	**Weight**	1
Domain	**Signalling question**			**Response**	**Comments**
Bias arising from the randomization process	1.1 Was the allocation sequence random?			Y	Quote: "Random assignment using a computer-generated table of random numbers was conducted by the project coordinator on the remaining 29 veterans."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?			Y	

			Quote from correspondence with the authour: "the allocation sequence for participants was blinded, so that only the project coordinator had access to the randomization table that the statistician prepared. The PI and other co-investigators were blinded."
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "[...] there were no significant differences on any demographic or outcome variables at baseline except for the number of years in the military. The mantram group reported an average of 3.9 (SD = 1.81) years compared to controls with 11.1 (SD = 9.90) years. Despite these differences, both groups had an equivalent number of months in combat averaging 11.2 ± 6.99 months."
	Risk of bias judgement	**Low**	computer-generated random allocation sequence; allocation sequence concealed from PI and co-investigators; baseline scores do not suggest a problem with randomization
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were necessarily aware of the assigned condition
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	Quote: "Efforts to maintain fidelity and quality of the intervention, included audiotaping randomly selected class sessions and having two outside intervention experts use a checklist to verify that the course content was taught."
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PN	Quote: "Retention was high suggesting that the intervention was acceptable."

			comment: accoring to report all randomized participants participated and completed treatment as planned; no non-adherence reported
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	reported information suggests that there were no substantial deviations from intended interventions
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	comment: 29/29 randomized participants were assessed at t2 and completed
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	pre- and post-assessment data available for all participants randomized
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: " To obtain objective assessments of PTSD symptom severity, research personnel were blinded to group assignment and conducted the ClinicianAdministered PTSD Scale (CAPS; Blake,et al.,1990)." comment: The administered scale (CAPS) is a validated PTSD measure and likely to be sensitive to intervention effects; it is considered gold-standard
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points

	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "To obtain objective assessments of PTSD symptom severity, research personnel were blinded to group assignment." comment: comment: although the assessors were blind to the participants' condition, there was no blind assessment: Accoring to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: although the interviewers were blinded, there is still a risk of bias. Because the comparator is a passive control condition the risk is higher that participants might have answered questions according to their beliefs/expectations regarding their assigned condition. This risk can not be fully eliminated by assessment by a blinded clinician.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. It is particularly high as the comparator was a no-treatment condition (WL)
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	
	5.3 ... multiple eligible analyses of the data?	PN	

	Risk of bias judgement		Some concerns	
Overall bias	**Risk of bias judgement**		**High**	comment: there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect (risk is especially relevant because comparator is a passive control condition). In addition, the assessor was no independent researcher who was not involved in the study, (which would otherwise lower the risk of bias). Interviews were administered by the research personnel.

Unique ID	15	**Study ID**	160101	**Assessor**	R
Ref or Label	Carletto 2016	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention ...**	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	EMDR	**Comparator**	REL	**Source**	Journal article(s) with results of the trial; Personal communication with trialist
Outcome	clinician-rated PTSD	**Results**	SMD(between)	**Weight**	1
Domain	**Signalling question**			**Response**	**Comments**
Bias arising from the	1.1 Was the allocation sequence random?			Y	Quote: "If they agreed, they signed the informed consent and they were randomized to the experimental group (EMDR) or to the control group (RT), using a block-wise randomization sequence
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?			NI	

randomization process			(block size of 10). The sequence was determined by an independent statistical consultant using the "Random Number Generators" function in SPSS version14.0."
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "Table 1 presents socio-demographic characteristics of these patients at baseline. There were no significant differences in demographics and in clinical characteristics between the two groups at baseline(T0)."
	Risk of bias judgement	**Some concerns**	a block-wise randomization sequence, NI on allocation concealment; no significant differences in demographics and in clinical characteristics.
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	NI	comment: therapists were necessarily aware of participants' assigned condition during treatment
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	comment: no non-protocol interventions reported
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	NI	
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PY	Quote: "Five patients did not begin the treatment (four in the EMDR group and one in the RT group) and three patients (one in the EMDR group and two in the RT group) attended only the first two sessions. These patients refused to continue with the assessment at T1 and T2 [...]"
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	PN	
	Risk of bias judgement	**High**	all in all, little information regarding deviations; participant non-adherence/dropout (20% in EMDR, 12% in REL); no appropriate analysis used to estimate the effect of adhering to the intervention.

Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	comment: EMDR: post-assessment data available for 20/25 participants randomized; REL: post-assessment data for 22 out of 25 participants randomized (additional data requested and subsequently provided by author)
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	comment: no information on such evidence in the report
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NI	comment: amount of missing data =16%; proportions of dropout approximately equal; no information on reasons for dropout reported.
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	
	Risk of bias judgement	**High**	20% of missing posttest data in EMDR group, 12% in REL group; no information on reasons for dropout reported.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	PN	Quote: "Then patients with a confirmed diagnosis of PTSD were assessed with the Clinician-Administered PTSD Scale (CAPS)"
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: " T2 was considered as the main assessment point throughout all the analyses. Assessments were independent and blind to treatment." comment: assuming that authors meant assessors by "assessment", assessors were blind to treatment condition However, although the assessors were blind to the

			participants' condition, there was no blind assessment: Accoring to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention. Thus, the risk that participants have answered according to their beliefs/expectations about the intervention effect is not as high (as opposed to a passive control condition)
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	N	
	Risk of bias judgement	**Some concerns**	unblinded assessors; the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: "The trial registrationnumber is NCT01743664." comment: Trial Registration and Changes of history raise no concerns.
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest; without leaving out any subscales
	5.3 ... multiple eligible analyses of the data?	PN	comment: raw values (means, SDs) (=unadjusted effect size estimates) are reported for all completers.
	Risk of bias judgement	**Low**	

Overall bias	Risk of bias judgement		High	a block-wise randomization sequence, NI on allocation concealment; no significant differences in demographics and in clinical characteristics. participant non-adherence/dropout; no appropriate analysis used to estimate the effect of adhering to the intervention. 20% of missing posttest data in EMDR group, 12% in REL group; no information on reasons for dropout reported. unblinded assessors; the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention.

Unique ID	30	Study ID	280103	Assessor	R
Ref or Label	Ehlers 2003	Aim	adhering to intervention (the 'per-protocol' effect)	The effect of adhering to intervention ...	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	CT	Comparator	WL	Source	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	"Assessors who decided whether patients were suitable for inclusion in the study could not predict what treatment condition would be assigned to the patient, as (1) the allocation list was kept locked in a seperate central office and the patient's allocation was only revealed 3 weeks later, following the self-monitoring assessment, and (2) the study was conducted at 2 sites."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	PY	

	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	education, education, profession, previous trauma, and comorbid depression); accident characteristics (eg, time since accident, injury severity, and severity of ongoing health problems); or initial symptom severity" "There were no site effects (Oxford vs Northampton) or interactions with site for any of the measures."
	Risk of bias judgement	**Low**	random permuted blocks with strata algorithm; allocation probably concealed; no group differences.
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	Quote: " they met with a clinician (other than the assessor) who informed them about their allocation and conducted the first session of the irrespective treatment condition."
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	Quote: "Each case was discussed in weekly group supervision meetings to ensure adherence to the treatment protocol." comment: the reported information does not indicate any failures in implementation; however, treatment fidelity was not systematically assessed
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PN	Quote: " although our aim was to recruit patients in the second month after trauma and to start intervention at 3 months, a few patients did not attend their initial appointments, making rescheduling necessary. This had the effect that interventions started on average at 4 months after the accident"

			comment: no other cases of non-adherence reported
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	no blinding of participants/therapists; no major deviations
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	see Figure 1 flow-chart comment: in CT all participants who were randomized completed treatment and were post-assessed. 2 out of 29 participants in the WL group were not post-assessed
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	no missing data
Bias in measurement	4.1 Was the method of measuring the outcome inappropriate?	N	comment: The administered scale (CAPS) is a validated, gold-standard PTSD measure and likely to be sensitive to intervention effects

| of the outcome | 4.2 Could measurement or ascertainment of the outcome have differed between intervention groups? | N | Quote: "Independent assessors (trained clinical psychologists or research nurses) who were not aware of the treatment condition gave the Clinician-Administered PTSD Scale (CAPS-SX) interview. A random sample of 38 CAPS interviews (8 different interviewers) was rated by a second clinician (7 different raters). Results indicated good reliability for the PTSD diagnosis (k=0.94) and total severity score (r=0.96)."

comment: in addition to the reliability testing as described in the quote, the same measurement methods and thresholds were used for all participants, and used at comparable time points |
| | 4.3 Were outcome assessors aware of the intervention received by study participants? | PY | Quote: "If patients were still suitable for the trial after the self-monitoring phase, they met with a clinician (other than the assessor) who informed them about their allocation [...]"

comment: although the assessors were blind to the participants' condition, there was no blind assessment: According to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary. |

	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	Quote: " Independent assessors (trained clinical psychologists or research nurses) who were not aware of the treatment condition gave the Clinician-Administered PTSD Scale (CAPS-SX) interview."
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NI	
			comment: the risk of bias due to knowledge of the intervention might be lowered by the fact that the outcome is not based solely on self-ratings of unblinded participants but reflects the clinical impression of improvement rated by an expert who was an independent researcher. It is, however, not clear from the reported information whether the independent assessors were also independent researchers (i.e. not involved in the trial). Thus, there is a risk of bias.
	Risk of bias judgement	**High**	risk of bias due to knowledge of the intervention because participants were not blinded and the comparator was a WL condition.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: both ITT and completers data is available; unadjusted as well as baseline adjusted effect sizes are reported; generally the risk of bias due to mutiple eligible analyses of the data is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Some concerns**	

Overall bias	Risk of bias judgement			High	comment: there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect (risk is especially relevant because comparator is a passive control condition).

Unique ID	42	Study ID	340101	Assessor	R
Ref or Label	Foa 1999	Aim	adhering to intervention (the 'per-protocol' effect)	The effect of adhering to intervention ...	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	PE	Comparator	SIT	Source	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	NI	Quote: " Those who met criteria for the study and signed consent forms were randomly assigned to one of the following four conditions: PE, SIT, combined treatment (PE-SIT), or WL. After enrolling 10 participants into WL, we assigned more participants to the three active groups than to WL."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
			comment: no information provided on the randomization procedure or allocation concealment. No reasons reported why fewer participants were allocated to the WL condition instead of using a 1:1

			allocation ratio. In the end, this imbalance in group sizes should not have a major impact on the results unless it is a result of a methodological bias in treatment allocation.
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	Quote: "Participants in the four treatment conditions did not differ significantly in their demographics and pretreatment measures of psychopathology, but there was a trend toward group differences on employment status, $x^2(3, N = 96) = 6.46, p = .09$. Nineteen percent of PE participants were nonworking compared with 30% of SIT, 43% of PE-SIT, and 8% of WL participants. No pre- or posttreatment differences were detected between victims of sexual and nonsexual assault."
	Risk of bias judgement	**Some concerns**	no information provided on the randomization procedure or allocation concealment
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	comment: therapists and participants were necessarily aware of the assigned condition
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	comment: no non-protocol interventions reported

| 2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome? | PN | Quote: " Therapists were trained to use manualsthat specified precise treatment guidelines for each session and received ongoing supervision by Edna B. Foa and Constance V. Dancu."

"Possible therapist effects were examined in a two-way (Therapist [4] < Condition [3]) analysis of covariance (ANCOVA) on the PSS-I, adjusting for pretreatment severity. Four therapists had each treated 5 or fewer participants. They were combined for these analyses and compared with the remaining three therapists. No significant main effects or an interaction were detected. Therapists also did not differ in dropout rate, $x^2(3, N = 81) = 4.48$, $p = .21$."

"Videotapes of 63 therapy sessions (9% of the 702 sessions) were randomly selected and rated. The adherence manual listed 52 treatment components that were present in any of the three protocols. Raters were familiar with the treatment programsbut had not treated any participants in this study. They reviewed videotapes and rated each componentas present or absent, without regard to treatment condition. On average, therapists completed 93% (SD = 12%) of the components prescribed for a given session in the corresponding protocol (PE, SIT, or PE-SIT). Only one deviation from the protocol was detected: 1 participant in the SIT protocol wasinstructed in the use of the Subjective Units of Distress scale, a componentprescribed in the PE and PE-SIT protocols. However, because this was not followed by exposure, this deviation was considered insignificant." |

	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	NI	comment: compliance was either not systematically assessed or it is nor reported. However, participants received the interventions they were assigned to (no crossovers to the comparator or switches to another active intervention are reported)
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	PN	
	Risk of bias judgement	**Some concerns**	no blinding of participants/therapists; no major deviations reported but adherence was not systematically assessed and there is insufficient information.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "Seventeen participants dropped out of treatment, leaving 79 completers. Dropouts were 2 (8%) of 25 PE participants, 7 (27%) of 26 SIT participants, 8 (27%) of 30 PE-SIT participants, and 0 of 15 WL participants." comment: 2 dropouts from PE, 8 from SIT
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no appropriate analysis correcting for bias or sensitivity analysis (or the like) reported
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	Quote: "The dropout rate differed significantly across groups, x7(3, N = 96) = 10.62, p < .025. More participants dropped out from SIT and PE-SIT (27%) than from the PE and WL conditions (5%), x'(1, N = 96) = 8.67, p < .01."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PY	"We did not conduct assessments during the active-treatment phase and therefore do not have information on the status of dropoutsat the time of termination. It is possible that some participants dropped out because they were doing well and were not motivated to complete the treatment.

			Because more participants dropped out from SIT and PE-SIT than from PE,this could have resulted in underestimating the efficacy of the former treatments."
			comment: In addition to the considerations quoted, the reason for drop-out could just as well have been the experience of worsening of symptoms or of adverse effects, thus overestimating the Treatment efficacy of SIT. The fact that authors did not consider this possibility in the report may even raise concerns
			All in all, missingness in the outcome could depend on its true value, as no reasons for dropout are reported. Since there might be a relation, proportions of dropout are examined. Significantly more participants dropped out from SIT than from PE. Therefore, the risk of bias is rated "high"
	Risk of bias judgement	**High**	missingness in the outcome could depend on its true value, as no reasons for dropout are reported. Significantly more participants dropped out from SIT than from PE.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	PN	Quote: "The PSS-I consists of 17 questions that correspond to the DSM-III-R PTSD symptoms, each rated on a 0-3-point scale for frequencyandseverity. Interrater reliability for both the diagnosis of PTSD (x = .91) and overall severity ratings (r = .97) are excellent (Foaet al., 1993)."

4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "we did not evaluate interrater reliability systematically for the PSS-I, the SAS ,and the SCID throughout the 5 years of this study, allowing for the possibility of rater drift. However, participants from the first 3 years of this study were all included in a psychometric study that demonstrated high reliability of the PSS-I (Foaet al., 1993)."
		comment: The quote indicates that the possibility of rater drift can not be ruled out completely. However, beyond that aspect there is no reason for concerns as the same measurement methods were used for all participants, and used at comparable time points
4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: " Independent evaluators were female clinicians with at least a master's degree who received extensive training in administration of the instruments and were unaware of treatment assignment."
		comment: although the assessors were blind to the participants' condition, there was no blind assessment: According to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.
4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the
4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN	comparator was also an active intervention -

				reducing the probability that participants have answered according to their beliefs/expectations about the intervention effect (in contrast to a passive control condition)
	Risk of bias judgement		Some concerns	no blind assessment; the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?		NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?		PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?		PN	comment: means and SDs are reported for completers. generally, the risk of bias due to selection from multiple eligible analyses of the data is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement		Some concerns	
Overall bias	**Risk of bias judgement**		High	insufficient or no information reported on randomization procedure, allocation concealment, reasons for dropout, reliability of interviewers' PTSD ratings, and pre-specified analysis plan. no blinding of participants/therapists; no major deviations reported but adherence was not systematically assessed and there is insufficient information. no blind assessment; the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention.

Unique ID	43	**Study ID**	340102	**Assessor**	R
Ref or Label	Foa 1999	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to**	occurance of non-protocol interventions; failures in implementing the intervention that could

				intervention ...	have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	PE	**Comparator**	WL	**Source**	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	NI	Quote: " Those who met criteria for the study and signed consent forms were randomly assigned to one of the following four conditions: PE, SIT, combined treatment (PE-SIT), or WL. After enrolling 10 participants into WL, we assigned more participants to the three active groups than to WL."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	comment: no information provided on the randomization procedure or allocation concealment. No reasons reported why fewer participants were allocated to the WL condition instead of using a 1:1 allocation ratio. In the end, this imbalance in group sizes should not have a major impact on the results unless it is a result of a methodological bias in treatment allocation.
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	Quote: "Participants in the four treatment conditions did not differ significantly in their demographics and pretreatment measures of psychopathology, but there was a trend toward group differences on employment status, $x^2(3, N = 96) = 6.46$, $p = .09$. Nineteen percent of PE participants were nonworking compared with 30% of SIT, 43% of PE-SIT, and 8% of WL participants. No pre- or posttreatment differences were detected between victims of sexual and nonsexual assault."

	Risk of bias judgement	Some concerns	no information provided on the randomization procedure or allocation concealment
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	comment: therapists and participants were necessarily aware of the assigned condition
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	comment: no non-protocol interventions reported
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	Quote: " Therapists were trained to use manualsthat specified precise treatment guidelines for each session and received ongoing supervision by Edna B. Foa and Constance V. Dancu." "Possible therapist effects were examined in a two-way (Therapist [4] < Condition [3]) analysis of covariance (ANCOVA) on the PSS-I, adjusting for pretreatment severity. Four therapists had each treated 5 or fewer participants. They were combined for these analyses and compared with the remaining three therapists. No significant main effects or an interaction were detected. Therapists also did not differ in dropout rate, $x^2(3, N = 81) = 4.48, p = .21$." "Videotapes of 63 therapy sessions (9% of the 702 sessions) were randomly selected and rated. The adherence manual listed 52 treatment components that were present in any of the three protocols. Raters were familiar with the treatment programsbut had not treated any participants in this study. They reviewed videotapes and rated each componentas present or absent, without regard to treatment condition. On average, therapists completed 93% (SD = 12%) of the components prescribed for a given session in the corresponding protocol (PE, SIT, or PE-SIT). Only one deviation from the

			protocol was detected: 1 participant in the SIT protocol wasinstructed in the use of the Subjective Units of Distress scale, a componentprescribed in the PE and PE-SIT protocols. However, because this was not followed by exposure, this deviation was considered insignificant."
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	NI	comment: compliance was either not systematically assessed or it is nor reported. However, participants received the interventions they were assigned to (no crossovers to the comparator or switches to another active intervention are reported)
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	PN	
	Risk of bias judgement	**Some concerns**	no blinding of participants/therapists; no major deviations reported but adherence was not systematically assessed and there is insufficient information.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	PY	Quote: "Seventeen participants dropped out of treatment, leaving 79 completers. Dropouts were 2 (8%) of 25 PE participants, 7 (27%) of 26 SIT participants, 8 (27%) of 30 PE-SIT participants, and 0 of 15 WL participants."

			comment: 2 dropouts (8%) from PE, 0 from WL
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	2 dropouts from PE, 0 from WL
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	PN	Quote: "The PSS-I consists of 17 questions that correspond to the DSM-III-R PTSD symptoms, each rated on a 0-3-point scale for frequencyandseverity. Interrater reliability for both the diagnosis of PTSD (x = .91) and overall severity ratings (r = .97) are excellent (Foaet al., 1993)."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "we did not evaluate interrater reliability systematically for the PSS-I, the SAS ,and the SCID throughout the 5 years of this study, allowing for the possibility of rater drift. However, participants from the first 3 years of this study were all included in a psychometric study that demonstrated high reliability of the PSS-I (Foaet al., 1993)." comment: The quote indicates that the possibility of rater drift can not be ruled out completely. However, beyond that aspect there is no reason for concerns as the same measurement methods were used for all participants, and used at comparable time points

	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: " Independent evaluators were female clinicians with at least a master's degree who received extensive training in administration of the instruments and were unaware of treatment assignment." comment: although the assessors were blind to the participants' condition, there was no blind assessment: According to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	comment: the risk of bias due to knowledge of the intervention is rated higher if the comparator is no treatment (here: WL) than when the comparator is another active intervention.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	there is a high risk of bias due to knowledge of the intervention because comparator is a passive control condition.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: means and SDs are reported for completers. generally, the risk of bias due to selection from multiple eligible analyses of the data is low as the results assessed here are raw values (means, SDs)

	Risk of bias judgement		Some concerns	
Overall bias	Risk of bias judgement		High	insufficient or no information reported on randomization procedure, allocation concealment, reliability of interviewers' PTSD ratings, and pre-specified analysis plan; in addition, there is a high risk of bias due to knowledge of the intervention because comparator is a passive control condition.

Unique ID	44	Study ID	340103	Assessor	R
Ref or Label	Foa 1999	Aim	adhering to intervention (the 'per-protocol' effect)	The effect of adhering to intervention …	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	SIT	Comparator	WL	Source	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	NI	Quote: " Those who met criteria for the study and signed consent forms were randomly assigned to one of the following four conditions: PE, SIT, combined treatment (PE-SIT), or WL. After enrolling 10 participants into WL, we assigned more participants to the three active groups than to WL." comment: no information provided on the randomization procedure or allocation concealment. No reasons reported why fewer participants were allocated to the WL condition instead of using a 1:1 allocation ratio. In the end, this imbalance in group
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	

			sizes should not have a major impact on the results unless it is a result of a methodological bias in treatment allocation.
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	Quote: "Participants in the four treatment conditions did not differ significantly in their demographics and pretreatment measures of psychopathology, but there was a trend toward group differences on employment status, x²(3, N = 96) = 6.46, p = .09. Nineteen percent of PE participants were nonworking compared with 30% of SIT, 43% of PE-SIT, and 8% of WL participants. No pre- or posttreatment differences were detected between victims of sexual and nonsexual assault."
	Risk of bias judgement	**Some concerns**	no information provided on the randomization procedure or allocation concealment
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	comment: therapists and participants were necessarily aware of the assigned condition
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	comment: no non-protocol interventions reported

| 2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome? | PN | Quote: " Therapists were trained to use manualsthat specified precise treatment guidelines for each session and received ongoing supervision by Edna B. Foa and Constance V. Dancu."

"Possible therapist effects were examined in a two-way (Therapist [4] < Condition [3]) analysis of covariance (ANCOVA) on the PSS-I, adjusting for pretreatment severity. Four therapists had each treated 5 or fewer participants. They were combined for these analyses and compared with the remaining three therapists. No significant main effects or an interaction were detected. Therapists also did not differ in dropout rate, $x^2(3, N = 81) = 4.48, p = .21$."

"Videotapes of 63 therapy sessions (9% of the 702 sessions) were randomly selected and rated. The adherence manual listed 52 treatment components that were present in any of the three protocols. Raters were familiar with the treatment programsbut had not treated any participants in this study. They reviewed videotapes and rated each componentas present or absent, without regard to treatment condition. On average, therapists completed 93% (SD = 12%) of the components prescribed for a given session in the corresponding protocol (PE, SIT, or PE-SIT). Only one deviation from the protocol was detected: 1 participant in the SIT protocol wasinstructed in the use of the Subjective Units of Distress scale, a componentprescribed in the PE and PE-SIT protocols. However, because this was not followed by exposure, this deviation was considered insignificant." |

	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	NI	comment: compliance was either not systematically assessed or it is nor reported. However, participants received the interventions they were assigned to (no crossovers to the comparator or switches to another active intervention are reported)
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	PN	
	Risk of bias judgement	**Some concerns**	no blinding of participants/therapists; no major deviations reported but adherence was not systematically assessed and there is insufficient information.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "Seventeen participants dropped out of treatment, leaving 79 completers. Dropouts were 2 (8%) of 25 PE participants, 7 (27%) of 26 SIT participants, 8 (27%) of 30 PE-SIT participants, and 0 of 15 WL participants." comment: 7 dropouts (27%) from SIT, 0 from WL
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis or the like reported
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	Quote: "The dropout rate differed significantly across groups, x7(3, N = 96) = 10.62, p < .025. More participants dropped out from SIT and PE-SIT (27%) than from the PE and WL conditions (5%), x'(1, N = 96) = 8.67, p < .01." "We did not conduct assessments during the active-treatment phase and therefore do not have information on the status of dropoutsat the time of termination. It is possible that some participants dropped out because they were doing well and were not motivated to complete the treatment.
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PY	

			Because more participants dropped out from SIT and PE-SIT than from PE, this could have resulted in underestimating the efficacy of the former treatments."
			comment: In addition to the considerations quoted, the reason for drop-out could just as well have been the experience of worsening of symptoms or of adverse effects, thus overestimating the treatment efficacy of SIT. The fact that authors did not discuss this possibility in the report may even raise concerns.
			All in all, missingness in the outcome could depend on its true value, as no reasons for dropout are reported. Since there might be a relation, proportions of dropout are examined: Significantly more participants dropped out from SIT than from WL. Therefore, the risk of bias is rated "high"
	Risk of bias judgement	**High**	missingness in the outcome could depend on its true value, as no reasons for dropout are reported. Significantly more participants dropped out from SIT than from WL.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	PN	Quote: "The PSS-I consists of 17 questions that correspond to the DSM-III-R PTSD symptoms, each rated on a 0-3-point scale for frequency and severity. Interrater reliability for both the diagnosis of PTSD (x = .91) and overall severity ratings (r = .97) are excellent (Foaet al., 1993)."

	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "we did not evaluate interrater reliability systematically for the PSS-I, the SAS ,and the SCID throughout the 5 years of this study, allowing for the possibility of rater drift. However, participants from the first 3 years of this study were all included in a psychometric study that demonstrated high reliability of the PSS-I (Foaet al., 1993)."
			comment: The quote indicates that the possibility of rater drift can not be ruled out completely. However, beyond that aspect there is no reason for concerns as the same measurement methods were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: " Independent evaluators were female clinicians with at least a master's degree who received extensive training in administration of the instruments and were unaware of treatment assignment."
			comment: although the assessors were blind to the participants' condition, there was no blind assessment: According to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	comment: the risk of bias due to knowledge of the intervention is rated higher if the comparator is no
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	

			treatment (here: WL) than when the comparator is another active intervention.
	Risk of bias judgement	**High**	there is a high risk of bias due to knowledge of the intervention because comparator is a passive control condition.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: means and SDs are reported for completers. generally, the risk of bias due to selection from multiple eligible analyses of the data is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Some concerns**	
Overall bias	**Risk of bias judgement**	**High**	insufficient or no information reported on randomization procedure, allocation concealment, reasons for dropout, reliability of interviewers' PTSD ratings, and pre-specified analysis plan; in addition, there is a high risk of bias due to knowledge of the intervention because comparator is a passive control condition.

Unique ID	27	**Study ID**	270101	**Assessor**	R
Ref or Label	Edmond 1999	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention ...**	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	EMDR	**Comparator**	WL	**Source**	Journal article(s) with results of the trial

Outcome	self-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question			Response	Comments
Bias arising from the randomizatio n process	1.1 Was the allocation sequence random?			NI	Quote: "The 59 participants were assigned randomly to one of three groups: (1) individual EMDR treatment (n = 20); (2) routine individual treatment (n = 20); or (3) delayed treatment control group (n = 19)."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?			NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?			PN	Quote: "No significant differences were found between groups on the basis of treatment or therapist assignment on any of the demographic characteristics, abuse specific variables, or pretest scores. Furthermore, no significant differences were found on pretests between survivors sexually abused by a single versus multiple perpetrators."
	Risk of bias judgement			**Some concerns**	no information on randomization procedure and allocation concealment provided
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?			Y	Quote: "Participants were not informed of their random treatment assignment until after completing the pretests"
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?			PY	
					comment: therapists and participants were probably necessarily aware of the assigned condition
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?			PY	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?			PN	Quote: "Although each session was expected to be held weekly so that each survivor could complete treatment within six weeks, scheduling conflicts resulted in some variations. There was a significant difference between groups on the mean length of time between pretesting and posttesting (10.4 weeks for the EMDR participants, 11 weeks for the routine individual treatment participants, and 7.4

			weeks for the control group). However, length of time was used as a covariate to assess for its effects on the posttest results and was found not to be a significant factor." comment: the reported variation in scheduling is not likely to substantially impact the outcome
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PN	comment: no non-adherence reported, no dropouts from either treatment condition.
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	no blinding of participants/therapists; no major deviations reported.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	Quote: "No attrition occurred during the pre- and posttest phase of the study" comment: results available for all 59 participants that were randomized
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	no attrition
Bias in measurement	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The IES is very sensitive to change and thus is viewed as an appropriate instrument for monitoring treatment progress (Corcoran & Fischer, 1987). The IES has shown groups validity and very

of the outcome			good internal consistency with alphas ranging from .79 to .92."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "There was a significant difference between groups on the mean length of time between pretesting and posttesting (10.4 weeks for the EMDR participants, [...] and 7.4 weeks for the control group). However, length of time was used as a covariate to assess for its effects on the posttest results and was found not to be a significant factor." comment: the same measurement methods and thresholds were used for all participants; the reported variation in treatment intervals and thus post-treatment assessment is not likely to have caused relevant differences in outcome the measurement
	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	comment: the IES is a self-report questionnaire; thus, the 'assessors' were -in this case- the participants themselves (who were aware of their intervention)
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	comment: could have been influenced because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. the risk of bias due to knowledge of the intervention is high, as the comparator is a passive control condition.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	risk of bias due to knowledge of the intervention is high because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect; it is particularly high as the comparator is a passive control condition.

Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?		NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?		NI	
	5.3 ... multiple eligible analyses of the data?		PN	comment: data for all participants randomized available; generally, the risk of bias due selection of results based on multiple eligible analyses is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement		Some concerns	
Overall bias	Risk of bias judgement		High	risk of bias due to knowledge of the intervention is high because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect; it is particularly high as the comparator is a passive control condition.

Unique ID	60	Study ID	410101	Assessor	R
Ref or Label	Ghafoori 2016	Aim	adhering to intervention (the 'per-protocol' effect)	The effect of adhering to intervention ...	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	Psychoeducation	Comparator	WL	Source	Journal article(s) with results of the trial
Outcome	self-rated PTSD	Results	SMD(between)	Weight	1
Domain	**Signalling question**			**Response**	**Comments**
Bias arising from the randomization process	1.1 Was the allocation sequence random?			Y	Quote: "Individuals who agreed to participate in the RCT portion of the study were randomized to receive a single session psychoeducation treatment or participate in a delayed treatment control group according a predetermined, computer-generated,
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?			NI	

			randomization list, which was developed by the statistician of the study (O.K.)."
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	NI	comment: baseline data and analyses of baseline characteristics only reported for completers (n=67), not for all participants randomized (n=86)
	Risk of bias judgement	**Some concerns**	no information regarding allocation concealment and results of analyses of baseline characteristics are only reported for the completer sample
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	Quote: " Session checklists were completed at the end of each session and monitored during weekly supervision to ensure the clinicians covered the components of psychoeducation. Review of session checklists indicated that all components of the psychoeducation treatment were completed with all participants in the study."
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PN	Quote: "One person randomized to the treatment condition withdrew in the middle of the treatment due to misunderstanding the amount of time the treatment would take."
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	no blinding, session checklists used to assess adherence - no major deviations;

Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "Of the 86 participants, 18 individuals who were randomized to the control arm of the study did not return for additional assessment. One person randomized to the treatment condition withdrew in the middle of the treatment due to misunderstanding the amount of time the treatment would take. Thus, 67 individuals completed both the pre (T0) and post (T1) assessments (37 psychoeducation treatment, 30 control)"
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis or the like reported
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	Quote: "One person randomized to the treatment condition withdrew in the middle of the treatment due to misunderstanding the amount of time the treatment would take."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PY	comment: difference regarding the proportions of missing outcome data at T2 between the groups (PsEd n=1 [2%]; WL n=18 [37.5%]). This leads to a risk of bias. In addition, no reasons for dropout from the WL condition are reported. One possible explanation may be that they were post-assessed after one month in contrast to the PsEd group that was post-assessed the same day that they were pre-assessed and received the treatment. It is also possible that they dropped out because they felt unlucky to be assigned to the waitlist condition. This would lead to bias as it is related to the assignes condition. Another possible explanation might be that symptoms remitted and those participants were no longer motivated to continue with the study. This would also lead to bias. In the end, the reasons remain unknown.

	Risk of bias judgement	High	difference regarding the proportions of missing outcome data at T2 between the groups (PsEd n=1 [2%]; WL n=18 [37.5%]). no reasons for dropouts from the WL group reported
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The PTSD Checklist-Civilian version (PCL-C; Weathers, Litz, Herman, Huska, & Keane, 1993) is a 17-item self-report PTSD symptom instrument that has been shown to have good internal consistency, strong correlations with other PTSD scales, and high diagnostic efficiency (Weathers et al., 1993)."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "First, considering the control group had their posttreatment assessment one month after the baseline assessment, whereas the treatment group had their post-treatment assessment the same day as their baseline assessment, [...]" comment: the same measurement methods and thresholds were used for all participants, but measurement time points (post-assessment) differed between groups. Errors in measuring of participants' outcome variables arise when the measured values do not equal the true or underlying values. Such errors can bias estimates of intervention effect from a randomized trial. This is not the case here.
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	comment: the PCL is a self-report questionnaire; thus, the 'assessors' were -in this case- the participants themselves (who were probably aware of their intervention)
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	

	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is rated higher because the comparator is a no-treatment condition (here: WL) (in contrast to the comparator being another active intervention).
	Risk of bias judgement	**High**	there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might answer according to their beliefs/expectations about the intervention effect
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	
	5.3 ... multiple eligible analyses of the data?	PN	comment: completers data are available; generally the risk of bias due to multiple eligible analyses of the data is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure of interest; for all time points of interest; completers data are available (not for the ITT sample); generally, the risk of bias due selection of results based on multiple eligible analyses is rather low considering that the results assessed here are raw values (means, SDs). All in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are some concerns.
Overall bias	**Risk of bias judgement**	**High**	no information allocation concealment, baseline characteristics/data only reported for the completer sample; risk of bias due to participants' knowledge of intervention; potential bias due to missing outcome data as reasons for dropout are not reported and proportions of missing outcome data

are different between groups; no information regarding pre-specified analysis plan.

Unique ID	61	Study ID	420101	Assessor	R
Ref or Label	Ghafoori 2017	Aim	adhering to intervention (the 'per-protocol' effect)	The effect of adhering to intervention ...	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	PE	Comparator	PCT	Source	Journal article(s) with results of the trial
Outcome	self-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "Participants who met eligibility criteria were then assigned to receive 12 sessions of PE treatment or 12 sessions of PCT treatment according to a predetermined, computer-generated randomization list. The randomization list was programmed to randomize two-thirds of participants to the PE treatment and one-third of the participants to the PCT treatment. The unequal allocation was justified ethically."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PY	Quote: "There were no statistically significant group differences with respect to demographic data at baseline (see Table 1) outside of race/ ethnicity." " Given that there was a significant mean difference between the PE and PCT groups at baseline for the PCL-5 and BSI-18-Depression measures, we controlled for this difference by including the baseline scores as a covariate in the model."

	Risk of bias judgement	High	unclear whether allocation sequence was concealed, significant difference between groups at baseline for the PTSD and Depression measures
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	Quote: ""Because of wide variation in therapist education and training, some inconsistency in delivery is likely and may impact the results." "Fidelity monitoring for the study from a two person review of audiotapes of treatment sessions revealed good adherence to the proscribed elements of the therapy protocols." comment: The quotes seem somewhat inconsistent. However, no specific failures of implementation are reported.
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PY	Quote: " Of the 24 individuals assigned to the PCT group, 4 completed less than 3 sessions, 20 completed 3 sessions, 18 completed 6 sessions, and 16 completed 9 sessions, and 8 completed 12 sessions." "Of the 47 individuals assigned to the PE group, 10 completed less than 3 sessions, 37 completed 3 sessions, 25 completed 6 sessions, 19 completed 9 sessions, and 13 completed 12 sessions." " No significant differences were found in the total number of sessions attended by treatment group, t(48.9) = 1.83, p≥ 0.05." "Fidelity monitoring for the study from a two person review of audiotapes of treatment sessions revealed good

			adherence to the proscribed elements of the therapy protocols. The raters judged 80% of the PE tapes to be satisfactory or better, defined as having adherence to all proscribed elements of the therapy session, and 100% of the PCT tapes tobe satisfactory or better with respect to adherence to the proscribed therapy elements."
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	PN	comment: An appropriate analysis would be inverse probability weighting to adjust for censoring of participants who cease adherence to their assigned intervention, in trials of sustained treatment strategies. This was not done.
	Risk of bias judgement	**High**	a large proportion of participants in both groups did not attend all 12 sessions. The reported information is not detailed and clear enough in order to reliably evaluate deviations from intended interventions and their potential impact on results. Also, the reasons for cessation are unknown. It can be assumed that reductions in PTSD symptom scores might have been larger if all participants had attended all 12 sessions. If, however, participants dropped out because of symptom worsening as a result of the treatment, then one would assume the opposite direction of bias. To conclude, it is difficult to predict the probability and direction of potential bias, raising concerns.

Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "Among the 16 individuals who did not complete the full 12 sessions of [PCT] treatment the reasons for dropout were no longer interested (n = 2), inconsistent attendance (n = 2), and unknown reasons (n = 20)."
			"The reasons for dropout [from PE] were scheduling conflict (n = 1), no longer interested (n = 2), childcare unavailable (n = 1), inconsistent attendance (n = 2), and unknown reasons (n = 41)."
			" Treatment dropout, defined as those who did not complete 12 sessions, was higher in the PE group (n = 34, 72%) than inthe PCT group (n = 16, 66%)."
			comment: see also Figure 1. The information provided is unclear concerning the number of participants that were post-assessed. The last quote above indicated that 16 participants dropped out from PE and 4 from PCT. However, details in the flow-chart and the first 2 quotes above list reasons for dropout for all 47 of the 47 participants assigned to PE and for all 24 of the 24 participants assigned to PCT. The amount of missing outcome data is unknown.
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	Quote: " The initial analysis consisted of data cleaning, which included examining the data for presence of outliers, conducting analysisof missing values, and constructing new variables."
			comment: neither are results of that analysis of missing data reported nor were analysis correcting

			for bias, sensitivity analysis or the like conducted (or not reported)
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	Quote: "Among the 16 individuals who did not complete the full 12 sessions of [PCT] treatment the reasons for dropout were no longer interested (n = 2), inconsistent attendance (n = 2), and unknown reasons (n = 20)."

"The reasons for dropout [from PE] were scheduling conflict (n = 1), no longer interested (n = 2), childcare unavailable (n = 1), inconsistent attendance (n = 2), and unknown reasons (n = 41)."

" Treatment dropout, defined as those who did not complete 12 sessions, was higher in the PE group (n = 34, 72%) than in the PCT group (n = 16, 66%)." "The highest reported reason for dropout was termed as unknown as contact ceased with the participants and follow-up was not possible."

comment: The amount of missing data is unknown, as are the proportions of missing outcome data in each group and reasons for dropout. If only those participants who completed all 12 sessions were post-assessed at that time point (after the regular 12 session period) - which is likely to be the case- then only 13 participants in the PE group and 8 in PCT were post-assessed. This would mean that 72.3% of the participants assigned to PE were not post-assessed, and 66.67% of participants |
| | 3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value? | NI | |

			assigned to PCT. The amount of missing data at t2 might be very cosiderable. The risk of bias is therefore rated high.
	Risk of bias judgement	**High**	The amount of missing data is unknown, as are the proportions of missing outcome data in each group and reasons for dropout. Therefore, the risk of bias is considered high.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The PCL-5 (Weathers et al., 2013) is a 20-item self-report PTSD symptom instrument that has been shown to have good internal consistency, strong correlations with other PTSD scales, and high diagnostic efficiency (Weathers et al., 1993)."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	comment: the PCL-5 is a self-report questionnaire; thus, the 'assessors' were -in this case- the participants themselves (who were probably aware of their intervention)
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention. Thus, the probability that participants have answered according to their beliefs/expectations about the intervention effect is not very high (as opposed to a passive control condition)
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN	
	Risk of bias judgement	**Some concerns**	participants were unblinded leading to a risk of bias (self-report PTSD measure). However, the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention.
Bias in selection of	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	

the reported result	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: see Table 3: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: raw values (means, SDs) (=unadjusted effect size estimates) are reported for the ITT sample (Baseline adjusted effect estimates (ANCOVA) were also reported). generally, the risk of bias due to multiple eligible analyses of the data is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure of interest for all time points of interest; details regarding missing data not reported; generally, the risk of bias due selection of results based on multiple eligible analyses is rather low considering that the results assessed here are raw values (means, SDs). All in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are some concerns.
Overall bias	**Risk of bias judgement**	**High**	no information regarding allocation concealment; concerns due to participants' knowledge of intervention; potentially high bias due to missing outcome data as the amount of missing data as well as reasons for dropout are unknown; no information regarding pre-specified analysis plan. SMD calculated from post-treatment scores might be biased due to between-group differences in baseline PTSD severity (PCT sign. lower than PE)

Unique ID	62	Study ID	500101	Assessor	R
Ref or Label	Hensel-Dittmann 2011	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to**	occurance of non-protocol interventions; failures in implementing the intervention that could have

				intervention ...	affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	NET	**Comparator**	SIT	**Source**	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomizatio n process	1.1 Was the allocation sequence random?	PY	Quote: " Subjects were randomly assigned to either NET or SIT. Participants were matched pairwise according to gender, age, and region of origin and were then allocated to NET or SIT by flipping a coin."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PY	Quote: " The participants of NET and SIT did not differ significantly in age, gender, education, years they had been living in Germany, comorbid psychiatric disorders, or area of origin." " We observed that NET and SIT differed in pretest mean PTSD scores. Although this difference did not reach statistical significance (possibly due to the small sample size) we cannot fully rule out an influence of pretreatment scores on treatment success with NET" comment: baseline differences not significant but assumingly due to small sample size as descriptevely the difference is large (NET M= 96.47(15.89), SIT M= 85.15 (12.95))
	Risk of bias judgement	**High**	allocation sequence generation was probably random; unclear allocation concealment; baseline differences not significant but assumingly due to

			small sample size as descriptevely the difference is large
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	Quote: " In 17 cases, we conducted the treatment with the aid of trained interpreters who did not know the patients beforehand. Patients requiring interpreters were equally distributed across the 2 treatment groups."
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	NI	
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PY	Quote: " Two patients in the NET group dropped out after the first session before any exposure treatment was carried out. As these patients were suicidal, outpatient treatment was no longer considered appropriate and they underwent inpatient psychiatric treatment instead. The third patient in the NET group dropped out after 4 sessions. The patient went into hiding from the police because of a realistic fear of deportation. The 2 patients who dropped out from SIT showed increasingly less treatment motivation and failed to attend sessions repeatedly so that treatment could not be completed" Quote: " ll sessions were videotaped and randomly analyzed in order to ensure treatment adherence. Treatment implementation was also discussed in team sessions."

			comment: therapists' adherence to treatment protocols was systematically assessed but results are not reported which raises some concerns
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	N	
	Risk of bias judgement	**High**	5 participants discontinued treatment; therapist adherence to protocols was assessed but results are not reported which raises some concerns; no analysis to estimate the effect of adherence
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	see figure 1 comment: n=15 assigned to NET, n=11 post-assessed [26.67% missing]; n=13 assigned to SIT, 10 of them post-assessed [23.07% missing])
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	comment: no analysis correcting for bias or sensitivity analysis (or the like) reported
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	Quote: " Two patients in the NET group dropped out after the first session before any exposure treatment was carried out. As these patients were suicidal, outpatient treatment was no longer
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	

			considered appropriate and they underwent inpatient psychiatric treatment instead. The third patient in the NET group dropped out after 4 sessions. The patient went into hiding from the police because of a realistic fear of deportation. The 2 patients who dropped out from SIT showed increasingly less treatment motivation and failed to attend sessions repeatedly so that treatment could not be completed" comment: proportions of missing outcome data did not differ; in two cases (SIT) the reasons for dropout might be related to the treatment ("increasingly less treatment motivation"). However, most dropouts are probably unrelated to treatment, so the risk of bias is not assumed to be very high
	Risk of bias judgement	**Some concerns**	some concerns because two dropouts might be related to the treatment condition; however, in light of this small proportion, the risk of bias is not judged to be very high. There are some concerns though
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: " e assessed sociodemographic data, experiences of organized violence using the vivo checklist of war, detention, and torture events [24] , PTSD diagnosis and severity with the Clinician-Administered PTSD Scale (CAPS)" comment: The CAPS is a validated measure and has been shown to be sensitive to intervention effects; it is considered gold standard
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	

4.3 Were outcome assessors aware of the intervention received by study participants?		Y	Quote: " We aimed to keep the assessors blind to the treatment conditions of the subjects; however, occasionally the treatment condition was revealed to the rater by responses from the patient." comment: not all interviewers were blinded. In addition, there was no blinded assessment as participants were probably aware of their assigned treatment condition
4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?		Y	comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention. Thus, the probability that participants have answered according to their beliefs/expectations about the intervention effect is not very high (as opposed to a passive control condition). Nevertheless, there is a risk of bias becuase interviewers might have known participants' treatment condition -For subjective outcomes such as 'clinical impression of improvement', knowledge of the intervention received could be highly influential.
4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?		PY	
Risk of bias judgement		**High**	interviewer blindness was broken by participants revealing details about intervention; participants were probably unblinded; the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention. Thus, the probability that participants have answered according to their beliefs/expectations about the intervention effect is not very high (as opposed to a passive control condition). Nevertheless, there is a risk of bias becuase interviewers might have known participants' treatment condition -For subjective outcomes such

			as 'clinical impression of improvement', knowledge of the intervention received could be highly influential. The primary investigators' research focusses on narrative exposure therapy, therefore, it is possible that interviewers were positively biased toward NET. The lack of blinding is considered to pose a risk of bias that might favour NET
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	NI	" e observed that NET and SIT differed in pretest mean PTSD scores. Although this difference did not reach statistical significance (possibly due to the small sample size) we cannot fully rule out an influence of pretreatment scores on treatment success with NET" comment: raw values (means, SDs) (=unadjusted effect size estimates) are reported for all completers. generally, the risk of bias due selection of results based on multiple eligible analyses is low as the results assessed here are raw values (means, SDs). no baseline adjusted results are reported. All in all, difficult to assess based on the information provided
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for all outcome measures of interest; for all time points of interest; completers data available; estimated means and SDs from ITT (mixed model) analysis not reported; only unadjusted effect sizes reported; the risk of bias due selection of results based on multiple

					eligible analyses is low as the results assessed here are raw values (means, SDs). All in all, difficult to assess due to lack of information; hence, there are some concerns
Overall bias	**Risk of bias judgement**			**High**	no information on allocation concealment; risk of bias due to participants' and interviewers' knowledge of intervention; risk of bias due to missing outcome data cannot be rueld out; no information regarding pre-specified analysis plan. baseline differences not significant but assumingly due to small sample size as descriptevely the difference is large so the effect size estimate (SMD) might be biased

Unique ID	65	Study ID	560101	Assessor	R
Ref or Label	Högberg 2007	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention ...**	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	EMDR	**Comparator**	WL	**Source**	Journal article(s) with results of the trial
Outcome	self-rated PTSD	**Results**	SMD(between)	**Weight**	1
Domain	**Signalling question**			**Response**	**Comments**
Bias arising from the	1.1 Was the allocation sequence random?			PY	Quote: "The randomization was done by picking a sealed ballot in the presence of a research nurse who coordinated the study and followed the participants through all phases."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?			NI	

randomization process	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	Quote: "Demographic data and trauma characteristics for the treatment group and control group are presented in Table 1. There were no statistically significant differences between groups" comment: no difference in pretreatment PTSD severity
	Risk of bias judgement	**Some concerns**	randomization was done by picking a sealed ballot; no information on allocation concealment; no baseline differences between groups
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PN	Quote: "Five treatments with all the recorded sessions were selected by a computerized chance-number program and evaluated based on a checklist provided by the EMDR Institute. All five treatments were judged as acceptable regarding fidelity to the EMDR protocol." "Three subjects dropped out after randomization but before treatment/WL. One of them had a strong aversive reaction to the SPECT examination and decided to interrupt the examination. Two other subjects left the study because of difficulty with finding time for the study."

			comment: three dropouts; however, they dropped out before treatment/WL; no "imperfect adherence" occurred in the remaining participants, no noncompliance, there were no cross-overs to the comparator or switches to another active intervention; the risk of bias due to non-adherence is considered low
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	no non-adherence or failure in implementation reported
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "We decided not to use an intent-to-treat (20) analysis because there were no drop-outs during EMDR. In fact, WL had more drop-outs than treatment group between randomization and start of treatment. In this case, intent-to-treat would have overestimated the treatment gains." comment: 3 dropouts (before treatment if radomized to EMDR); given the fact that the total sample size is very small, even three dropouts could potentially impact results
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PN	Quote: Three subjects dropped out after randomization but before treatment/WL. One of them had a strong aversive reaction to the SPECT examination and decided to interrupt the examination. Two other subjects left the study
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	

			because of difficulty with finding time for the study. "We decided not to use an intent-to-treat (20) analysis because there were no drop-outs during EMDR. In fact, WL had more drop-outs than treatment group between randomization and start of treatment. In this case, intent-to-treat would have overestimated the treatment gains." comment: only one dropout where the reason for dropout was a study-related adverse event; however, the adverse reaction was not related to the treatment condition but to the SPECT (measurement of a physiological parameter)
	Risk of bias judgement	**Low**	only one dropout where the reason for dropout was a study-related adverse event; however, the adverse reaction was not related to the treatment condition but to the SPECT (measurement of a physiological parameter)
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "Self-rating scales were administered immediately before interviews 1 and 2. The assessments were performed on a computer with the research nurse present in the room. The Impact of Event Scale (IES) is a 15-item report instrument to evaluate the actual level of stress related to a past stressful event" comment: The administered scale (IES) is a validated PTSD measure and likely to be sensitive to intervention effects

	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points. see also figure 2
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	comment: the IES is a self-report questionnaire; thus, the 'assessors' were -in this case- the participants themselves (who were probably aware of their intervention)
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. It is particularly high as the comparator was a no-treatment condition (WL)
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. It is particularly high as the comparator was a no-treatment condition (WL)
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	
	5.3 ... multiple eligible analyses of the data?	PN	comment: , the risk of bias due selection of results based on multiple eligible analyses is rather low considering that the results assessed here are raw values (means, SDs).
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure of interest; for all time points of interest; completers data are available (not for the ITT sample); generally, the risk of bias due selection of results based on multiple eligible analyses is rather low considering that the results assessed here are

			raw values (means, SDs). All in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are some concerns.
Overall bias	**Risk of bias judgement**	**High**	no information on allocation concealment; risk of bias due to participants' knowledge of intervention; potential bias due to missing outcome data as reasons for dropout are not reported in sufficient detail; no information regarding pre-specified analysis plan

Unique ID	66	**Study ID**	630101	**Assessor**	R
Ref or Label	Ironson 2002	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention ...**	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	PE	**Comparator**	EMDR	**Source**	Journal article(s) with results of the trial
Outcome	self-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	PN	Quote: "After 20 patients had been randomly assigned to PE or EMDR, we noted more dropout in PE. To get a reasonably equivalent number of completers in both groups, the next two patients were assigned to PE."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	

	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	NI	Quote: " a perusal of initial scores suggested a difference between those assigned to PE and EMDR [In fact, initial scores on the PTSD scale were not significantly different between the groups, t(17) 5 1.76, p 5 .10; however, scores on the BDI were significantly higher among those assigned to PE, t(17) 5 2.23, p 5 .05.]" comment: apart from BDI and PSS-SR scores, no information regarding potential baseline characteristics of participants was provided.
	Risk of bias judgement	**High**	given the fact that randomization was broken and there is no information regarding allocation concealment and little information on baseline characteristics of participants, the risk of bias is considered high
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as blinding is not possible in the context of psychlogical interventions Quote: "Another limitation that might be considered is that the assessors who administered the measures were not blind to treatment condition."
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	NI	Quote: "Therapist variables contributing to the need for more than three EMDR sessions may include lack of the use of cognitive interweave (CI). Although the supervisors were trained in this EMDR technique, several clinicians were not trained in

			Level II and may not have had sufficient experience with CI, contributing to the difficulty with some of the EMDR cases."
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PY	Quote: "As hypothesized, dropout after the first active session was significantly higher for those randomized to treatment with PE (3 of 10 for PE, 0 of 10 for EMDR" " A further limitation is that although fidelity of the therapists was monitored by the supervisors listening to tapes, it was not systematically assessed." comment: according to Table 2 there were n=12 participants assigned to PE, six of them dropped out; and there were n=10 participants assigned to EMDR, none of them dropped out. The quote above is inconsistent with the numbers from the table. anyways, there were six dropouts from PE - three during sessions 1-3 and three during sessions 4-6. The former three were excluded from analysis
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	PN	comment: naive 'per protocol' analysis used; no inverse probability weighting to adjust for censoring of participants who cease adherence to their assigned intervention
	Risk of bias judgement	**High**	therapists and participants were probably necessarily aware of the assigned condition, six dropouts from PE treatment (3 druing sessions 1-3, 3 during sessions 4-6), treatment fidelity and therapists' adherence to protocols was not systematically assessed

Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "As hypothesized, dropout after the first active session was significantly higher for those randomized to treatment with PE (3 of 10 for PE, 0 of 10 for EMDR" " Thus, with PE we know with certainty at least four (and possibly another three who left before evaluation of the sixth active session) of the 12 enrolled met criterion after being offered [...]" comment: according to Table 2 there were n=12 participants assigned to PE, six of them dropped out; and there were n=10 participants assigned to EMDR, none of them dropped out. The quote above is inconsistent with the numbers from the table. anyways, there were six dropouts from PE - three during sessions 1-3 and three during sessions 4-6. The former three were excluded from analysis. the three participants who dropped out after session 3 did not complete post-assessments. data of 27.27% of participants that were randomized would be missing
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	Quote: "As noted, dropout was only considered attributable to a treatment if it occurred after one active session (either PE or EMDR), which meant that participants had completed the three preparatory sessions [...]" comment: The authors reason that reasons for
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	

			dropout before session 4 (n=3) cannot be related to either treatment condition (PE, EMDR) as the first 3 sessions were equivalent (regarding the content) for both conditions. Although this fact does lower the risk of bias, it might still be possible that dropout was related to the assigned condition: Given the fact that participants were probably aware of their assigned tratment condition, it is still possible that beliefs/expectations/interpretations regarding this treatment condition influenced their decision to drop out. It is also not clear if participants were informed that the first three sessions would be preparatory sessions and that sessions 4-6 would be treatment specific. Such information could have motivated participants to drop out before session 4 to avoid PE treatment.

Furthermore, even if only the remaining three dropouts are considered, the proportions of dropout are unequal and reasons for dropout are not reported. The risk of bias is therefore high |
| | **Risk of bias judgement** | High | unequal proportions of dropout, no reasons reported |
| **Bias in measurement of the outcome** | 4.1 Was the method of measuring the outcome inappropriate? | N | Quote: " PSS-SR items correspond to DSMIII-R diagnostic criteria and specifically address the symptom clusters of re-experiencing, avoidance/numbing, and arousal. Cronbach's alpha has been reported as .91 for total PSS-SR scores, and .78, .80, and .82 for the re-experiencing, avoidance, and arousal scales, respectively. One-month test-retest reliability ranged from .74 for total scores and .66, .56, and .71 for re-experiencing, avoidance, and arousal, respectively. The PSS-SR was validated against the Structured Clinical Interview for DSM-IIIR Disorders (SCID), and found |

			to correctly identify 86% of those with a SCID diagnosis of PTSD (Foa et al., 1993). In addition, the PSS-SR demonstrated concurrent validity with other measures of pathology; the BDI, r(42) 5 .80, p , .001, and the Rape Aftermath Symptom Test, r(42) 5 .81, p , .001)."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "Another limitation that might be considered is that the assessors who administered the measures were not blind to treatment condition." comment: therapists and participants were probably necessarily aware of the assigned condition as blinding is not possible in the context of psychlogical interventions comment: the PSS-SR is a self-report questionnaire; thus, the 'assessors' were -in this case- the participants themselves (who were probably aware of their intervention)
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention. Thus, the probability that participants have answered according to their beliefs/expectations about the intervention effect is not very high (as opposed to a passive control condition). However, assessors were not blinded and it is possible that they interacted with participants during post-assessment,
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NI	

			even though questionnaires were self-report instruments
	Risk of bias judgement	**High**	participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect, risk of bias is lowered by the fact that the comparator is also an active intervention. Assessors were not blind and it is unclear whether interactions took place during post-assessment
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	
	5.3 ... multiple eligible analyses of the data?	PN	comment: the risk of bias due selection of results based on multiple eligible analyses is relatively low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure of interest; for all time points of interest, but not regarding basline differences in demographic characteristics of participants; completers data reported; generally, the risk of bias due selection of results based on multiple eligible analyses is rather low considering that the results assessed here are raw values (means, SDs). All in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are concerns.

Overall bias	Risk of bias judgement		High	randomization was broken and there is no information regarding allocation concealment; unequal proportions of dropout, no reasons reported; participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect, risk of bias is lowered by the fact that the comparator is also an active intervention. Assessors were not blind and it is unclear whether interactions took place during post-assessment; no information regarding pre-specified analysis plan

Unique ID	68	Study ID	680101	Assessor	R
Ref or Label	Jarero 2015	Aim	adhering to intervention (the 'per-protocol' effect)	The effect of adhering to intervention ...	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	EMDR	Comparator	WL	Source	Journal article(s) with results of the trial
Outcome	self-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	NI	Quote: "During this time, two qualified, not blind to the research protocol, independent assessors explained the purpose of the research, as well as inclusion and exclusion criteria, obtained the informed consents, collected the clinical history of each participant, and applied the SPRINT as a baseline assessment for all participants (Time 1; Figure 1). During this phase, participants were divided randomly into two groups (immediate treatment condition and waitlist/delayed treatment condition) and randomly assigned to the three therapists."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	

			comment: no further details regarding randomization procedure reported
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	NI	comment: no information on baseline demographic characteristics. no significant difference of pretreatment PTSD scores
	Risk of bias judgement	**Some concerns**	no information on randomization or baseline demographic characteristics, but no significant difference of pretreatment PTSD scores
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition (as it is hardly possible to achieve blinding when psychological interventions are implemented)
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	Quote: "Treatment fidelity was fulfilled by strict observance to all steps of the scripted EMDR PRECI"
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PN	Quote: "Participation was voluntary, and there were no dropouts throughout the study period."

comment: no non-adherence reported |
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	

	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	NI	Quote: " The SPRINT (Connor & Davidson, 2001; Vaishnavi et al., 2006) is an 8-item interview or self-rating questionnaire with solid psychometric properties that can serve as a reliable, valid, and homogeneous measurement of PTSD illness severity and global improvement as well as a measure of somatic distress; stress coping; and work, family, and social impairment. [...] It was found that in the SPRINT, a cutoff score of 14 or more carried out a 95% sensitivity to detect PTSD and 96% specificity for ruling out the diagnosis, with an overall accuracy of correct assignment being 96% (Connor & Davidson, 2001)."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "On March 25, 2015, two independent assessors applied the SPRINT to participants in both groups (Time 2). " comment: (1) no blinding of interviewers reported; (2) even if the assessors were blind to the participants' condition, there was no blind assessment: According to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.

	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	comment: there is a risk of bias because participants were probably aware of their assigned condition and might have answered according to their beliefs/expectations regarding treatment efficacy. the risk of bias due to knowledge of the intervention is rated higher because the comparator is a no-treatment condition (here: WL) (in contrast to the comparator being another active intervention). In addition, the interviewer might have known the participants' treatment condition -For subjective outcomes such as 'clinical impression of improvement', knowledge of the intervention received could be highly influential.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	measure "SPRINT" is not very widely used and has only 8 items so there are concerns about its sensitivity and specifity; high risk of bias because no blinded assessment and a passive control condition.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	comment: no information provided; no trial registration
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	
	5.3 ... multiple eligible analyses of the data?	PN	comment: data for all participants randomized available; generally, the risk of bias due selection of results based on multiple eligible analyses is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Some concerns**	no information regarding a pre-specified analysis plan; the risk that the numerical result was selected on the basis of the results from multiple eligible analyses of the data is relatively low as the results considered there are raw values (means, SDs). All

			in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are some concerns.
Overall bias	**Risk of bias judgement**	**High**	no information on randomization procedure and allocation concealment; measure "SPRINT" is not very widely used and has only 8 items so there are concerns about its sensitivity and specifity; high risk of bias because no blinded assessment and a passive control condition. no information regarding pre-specified analysis plan

Unique ID	69	**Study ID**	690101	**Assessor**	R
Ref or Label	Jensen 1994	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention ...**	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	EMDR	**Comparator**	WL	**Source**	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	**Signalling question**	**Response**	**Comments**
Bias arising from the randomization process	1.1 Was the allocation sequence random?	NI	Quote: "Following the screening/initial assessment, the subject was randomly as- signed to either the EMD/R (N = 13) or control (N = 12) condition."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
			comment: no details provided regarding the randomization procedure and no information on allocation concealment

	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	NI	comment: no significance test of differences on demographic or clinical variables at baseline reported; demographic characteristics not provided for the intervention groups; descriptive group differences in PTSD symptom severity at baseline (EMDR M= 29.92; WL M= 37.08) (no significance test done)
	Risk of bias judgement	**Some concerns**	no information provided regarding random sequence generation, allocation concealment or baseline differences between intervention groups
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	Quote: "due to wording in the consent form, control subjects knew, after random assignment, that they were not in the experimental group."
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NI	Quote: " Control subjects were not deterred from pursing other mental-health treatment during the study and were given a list of local alternative treatment sites upon their being informed of their control group status." "The actual effects of these contacts, however, were likely in- significant because (1) all concurrent therapeutic contacts were not focused on the specific symptomatology targeted by EMD/R, and (2) it is unlikely that any of the control subjects were able to locate and obtain actual therapeutic benefit from any of the referral sources within the 17 days (approximately) from initial interview to final treatment session. This would seem especially unlikely in consideration of the fact that many of them had been seeking and falling to receive help for approximately 20 years."

			comment: it was not systematically assessed whether participants in the WL group received therapy elsewhere during the study.
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PY	Quote: "As previously indicated, one independent rater, expe- rienced in providing EMD/R treatment, did observe several videotaped treat- ment and history-taking sessions. Commenting on the observed level of treat- ment integrity, this individual made the following conclusion: "Conclusions based on EMD/R presented in this way could, theoretically, be somewhat sup- portive of EMD/R as a therapeutic modality. However, negative results could not be used to criticize EMD/R, as the clients may have received enough treat- ment to open difficult areas, but without enough fidelity to the treatment to resolve these problems" (H. J. Lipke, personal communication, October 25, 1991). The primary flaw, in this rater's opinion, was that the therapists did not appear to stay with, or continue in the active treatment phase with the videotaped subjects long enough to achieve resolution of symptoms"
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	NI	comment: no non-adherence reported beyond the four dropouts. It is not reported which intervention group those 4 had been assigned to.

	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	PN	comment: nothing of the sort reported
	Risk of bias judgement	**High**	no blinding of participants or therapists delivering th intervention; little information regarding deviations from intended interventions, making it difficult to assess the risk of bias
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "Twenty-nine subjects were eventually selected and volunteered for the study, with 4 choosing not to complete the study." comment: 13,8% dropped out and provided no posttest data (only completers are analysed)
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	comment: no sensitivity analyses or the like were conducted
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NI	comment: no information provided regarding proportions of dropout in the two intervention groups (it is unclear which group the 4 dropouts had been assigned to); no information regarding reasons for dropout
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	
	Risk of bias judgement	**High**	13.8% dropout (missing posttest data); no information provided regarding proportions of dropout in the two intervention groups (it is unclear which group the 4 dropouts had been assigned to); no information regarding reasons for dropout
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	PN	Quote: "Davidson et al. (1989) established the SI-PTSD's construct validity through factor analysis of data obtained from 116 interviewed veterans. They estab- lished content validity by producing high inter-correlations between subjects' total SI-PTSD scores and scores on a depression scale, an anxiety scale, and a PTSD measure. These authors established concurrent validity by cross- validating

			the SI-PTSD's diagnostic sensitivity and specificity with that of an- other commonly used structured clinical interview for PTSD."
4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN		
4.3 Were outcome assessors aware of the intervention received by study participants?	Y		Quote: " due to wording in the consent form, control subjects knew, after random assignment, that they were not in the experimental group." comment: it is not reported whether interviewers were blinded which raises concerns that they were not. Regardless of the blinding of interviewers, however, there was no blind assessment: According to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.
4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y		Quote: "Aware that they were not receiving treatment, these subjects could have had differential expectations regarding symptom improvement, which could have affected their posttest scores." comment: the risk of bias due to knowledge of the intervention is rated higher because the comparator
4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NI		

			is a no-treatment condition (here: WL) (in contrast to the comparator being another active intervention). In addition, the interviewer might have known the participants' treatment condition - For subjective outcomes such as 'clinical impression of improvement', knowledge of the intervention received could be highly influential.
	Risk of bias judgement	**High**	high risk of bias because participants were aware of their treatment condition and it is unclear whether interviewers were blinded. The risk is particularly hight because the comparator is a no-treatment condition (here: WL). In addition, the interviewer might have known the participants' treatment condition - For subjective outcomes such as 'clinical impression of improvement', knowledge of the intervention received could be highly influential.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	
	5.3 ... multiple eligible analyses of the data?	PN	comment: the risk of bias due selection of results based on multiple eligible analyses is relatively low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure of interest, for all time points of interest; completers data are available (not for the ITT sample); the risk of bias due selection of results based on multiple eligible analyses is relatively low considering that the results assessed here are raw values. All in all, however, it is difficult to assess the

			risk of bias in this domain due to lack of information; There are concerns.
Overall bias	**Risk of bias judgement**	**High**	no information provided regarding random sequence generation, allocation concealment or baseline differences between intervention groups; no blinding of participants or therapists delivering th intervention; little information regarding deviations from intended interventions, making it difficult to assess the risk of bias; 13.8% dropout (missing posttest data) but no information provided regarding proportions of dropout in the two groups or regarding reasons for dropout; high risk of bias because participants were aware of their treatment condition and it is unclear whether interviewers were blinded; no information regarding pre-specified analysis plan

Unique ID	70	**Study ID**	700101	**Assessor**	R
Ref or Label	Jindani 2015	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention ...**	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	MBI	**Comparator**	WL	**Source**	Journal article(s) with results of the trial
Outcome	self-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	**Signalling question**	**Response**	**Comments**
	1.1 Was the allocation sequence random?	Y	

	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	Quote: "Participants meeting entry criteria were assigned by a random number generator to either the experimental (yoga) group or waitlist control."
Bias arising from the randomizatio n process	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	Quote: "The only significant difference between the two groups was that males made up a slightly larger portion of the waitlist group before test. All of the males who began the study completed it. The two groups were compared at baseline on all outcome variables using independent samples t-tests. At baseline,the only scale on which the differences between study groups reached statistical significance was the PCL-17 scores (t(78) = 2.39, p = 0.019). Mean scores (M) demonstrated that the intervention group had higher PTSD scores (PCL; M = 59.48, SD = 9.33) at baseline than the waitlist control group (M= 55.14, SD= 11.86). " [comment: this analysis included all participants randomized]
	Risk of bias judgement	**Some concerns**	random allocation sequence; unclear allocation concealment; intervention group had higher PTSD scores at baseline than WL
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	Quote: " Fourth, although the participants were randomized to groups it was not possible to blind participants"
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	comment: therapists were probably necessarily aware of the assigned condition, too, as blinding is not possible when psychological interventions are implemented
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	Y	Quote: "Approximately 57% of the waitlist control group sought alternative treatment while 39% of the yoga group was involved in other therapies. This difference was nonsignificant, χ2= 2.8, NS."

	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	NI	
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	Y	Quote: "Ten individuals did not begin the study due to scheduling conflicts, 8 participants were not able to complete the program due to medical and health reasons, 4 participants discontinued for personal reasons, and 8 participants had schedule changes, missed classes, or were on vacations eventually leading to study dropout" comment: 20 participants in the Yoga group withdrew during treatment
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	PN	comment: naive 'per protocol' analysis used; no inverse probability weighting to adjust for censoring of participants who cease adherence to their assigned intervention
	Risk of bias judgement	**High**	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "This resulted in a 30% dropout rate with 29 participants completing the yoga program and 21 in the waitlist control group."
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis reported
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	Quote: "Ten individuals did not begin the study due to scheduling conflicts, 8 participants were not able to complete the program due to medical and health reasons, 4 participants discontinued for personal reasons, and 8 participants had schedule changes, missed classes, or were on vacations eventually leading to study dropout" "Second, the attrition in the yoga group likely
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	

			affected the results. It is not known how much of the results were due to the differential dropout rate of the yoga versus waitlist group." comment: there is a risk of bias due to (1) differential dropout rates between groups, and (2) some reasons for dropout (e.g. "health reasons"; "missed classes") might be related to the treatment
	Risk of bias judgement	**High**	there is a risk of bias due to (1) differential dropout rates between groups, and (2) some reasons for dropout (e.g. "health reasons"; "missed classes") might be related to the treatment
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The Posttraumatic Stress Disorder Checklist (PCL-17) is a validated 17-item self-report scale [24]. Cronbach's alpha has ranged from 0.94 [24] to 0.97 [25], and the test-retest reliability was 0.96 at 2-3 days and 0.88 at 1 week [24]." comment: The administered scale (PCL) is a validated PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	comment: the PCL is a self-report questionnaire; thus, the 'assessors' were -in this case- the participants themselves (who were aware of their intervention)
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	

	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	comment: there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. It is particularly high as the comparator was a no-treatment condition (WL)
	Risk of bias judgement	**High**	risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. It is particularly high as the comparator was a no-treatment condition
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	comment: no information provided
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	
	5.3 ... multiple eligible analyses of the data?	PN	comment: raw values (means, SDs) (=unadjusted effect size estimates) are reported for all completers. (Baseline adjusted effect estimated (ANCOVA) were also reported). generally, the risk of bias due to multiple eligible analyses of the data is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure of interest; for all time points of interest; completers data are available (not for the ITT sample); generally, the risk of bias due selection of results based on multiple eligible analyses is rather low considering that the results assessed here are raw values (means, SDs). All in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are some concerns.

Overall bias	Risk of bias judgement		High	no information on allocation concealment; risk of bias due to participants' knowledge of intervention; potential bias due to missing outcome data as there were differential dropout rates between groups and some of the reasons for dropout might have been related to the treatment; no information regarding pre-specified analysis plan. MBI group had higher PTSD scores at baseline than WL so SMD calculated from posttreatment scores might be biased

Unique ID	71	Study ID	720101	Assessor	R
Ref or Label	Kelly 2016	Aim	adhering to intervention (the 'per-protocol' effect)	The effect of adhering to intervention ...	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	MBI	Comparator	WL	Source	Journal article(s) with results of the trial
Outcome	self-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	NI	Quote: "A randomized waitlist controlled design was implemented to assess preliminary efficacy of the TI-MBSR intervention. Participants meeting eligibility criteria completed the informed consent process as well as the pre-intervention questionnaires, and were then randomly allocated to 8 weeks of TI-MBSR or a waitlist control condition (allocation ratio 1:1)." "Of the 45 women, 24 were randomly assigned to TI-MBSR and 21 to a wait-list control condition. Participants were informed of group allocation by phone."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	

	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "No statistically significant, between-groups differences were found at pre-intervention assessment for age, gender, race, income, exposure to traumatic violence, psychiatric symptoms, or attachment style."
	Risk of bias judgement	**Some concerns**	no information provided regarding random sequence generation and allocation concealment. No significant baseline differences between groups on relevant variables.
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	Quote: "Participants were informed of group allocation by phone."
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	comment: therapists were probably necessarily aware of the assigned condition
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NI	Quote: "Additionally, participants might have been in concurrent psychotherapy outside of the study, which may have influenced outcomes."
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	NI	Quote: " Also, the study lacked quantitative measurement of treatment fidelity: therapist adherence and competence was not evaluated."
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	Y	Quote: "The average number of participants per group session was 17/19. For those that attended at least the first session in the intervention group, 53% (n = 10) completed all eight sessions of the intervention, 26% (n = 5) completed 7 sessions, 11% (n = 2) completed 6 sessions, and 5% (n = 1) completed 5 sessions. One participant attended only the first class and did not return."
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	N	comment: no appropriate analysis used, such as inverse probability weighting to adjust for censoring of participants who cease adherence to their assigned intervention

	Risk of bias judgement	High	participants and therapists not blinded; authors state that participants might have received psychotherapy outside of the study; no assessment of treatment fidelity; 5 participants in the MBI group never started treatment and 9 did not attend all sessions. No appropriate analysis was used to estimate the effect of adhering to the intervention
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "Of the 45 women, 24 were randomly assigned to TI-MBSR and 21 to a wait-list control condition." "The number of women completing surveys at both pre- and post-intervention were 19 for the intervention group and 20 for the wait-list control group."
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	Quote: "To analyze patterns of missing data, we performed Little's MCAR test (Little, 1988). The pattern of missing data was consistent with being missing completely at random." comment: mechanism underlying missingness assessed but no sensitivity analyses conducted
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NI	Quote: " Approximately 87% (n=39) of the total enrolled sample (N=45) completed post-intervention assessments, with 20 participants completing the TIMBSR intervention and 19 completing the post-intervention measures from the wait-list control group; attrition did not significantly differ between groups, nor was it predicted by baseline values of study variables."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	

			comment: no sign. difference in proportions of missingness between groups. Little's MCAR test indicated that data is missing at random. no reasons for dropout reported.
	Risk of bias judgement	**Some concerns**	missing posttest data for 4.8% of participants in WL group and for 24% in MBI group; difference not significant. Authors conducted Little's MCAR test: results indicate that data is missing at random. Baseline values of study variables not predictive of attrition. No sensitivity analysis. No reasons for dropout reported.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "—The total score of the PTSD Checklist-Civilian Version (PCL-C) (Weathers, Litz, Herman, Huska, & Keane, 1994) was used to assess PTSD symptoms for participants. This 17-item self-report measure was selected due to its excellent reliability, validity, and generalizability to a wide variety of populations (Elhai, Gray, Kashdan, & Franklin, 2005; Elhai, Gray, Docherty, Kashdan, & Kose, 2007)."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	Quote. "All therapeutic change was measured via self-report, which may be biased and may not accurately measure clinical outcomes." comment: the PCL is a self-report questionnaire. therefore, the 'assessors' were the participants themselves - who were aware of their intervention
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	

	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	comment: assessment could have been influenced because participants were not blind to their assigned condition and might have answered according to their beliefs/expectations about the intervention effect comment: the risk of bias is particularly high as the comparator was a no-treatment condition (WL).
	Risk of bias judgement	**High**	comment: there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. It is particularly high as the comparator was a passive control condition
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	NI	comment: means and SDs reported for completers, but not for the ITT sample. generally, the risk of bias due to multiple eligible analyses of the data is loweredby the fact that results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Some concerns**	no pre-specified analysis plan available; no trial registration in non-commercial trial registry record; results are reported for the outcome measure of interest; for all time points of interest; completers data available (not for the ITT sample); generally, the risk of bias due selection of results based on multiple eligible analyses is relatively low considering that the results assessed here are raw values (means, SDs). All in all, it is difficult to

					assess the risk of bias in this domain due to lack of information; There are some concerns.
Overall bias	**Risk of bias judgement**			**High**	no information on randomization procedure and allocation concealment; risk of bias due to participants' knowledge of intervention; potential bias due to missing outcome data as reasons for dropout are not reported and no sensitivity analyses were conducted; no information regarding pre-specified analysis plan

Unique ID	72	**Study ID**	730101	**Assessor**	R
Ref or Label	Kim 2013	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention ...**	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	MBI	**Comparator**	WL	**Source**	Journal article(s) with results of the trial
Outcome	self-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	**Signalling question**	**Response**	**Comments**
Bias arising from the randomization process	1.1 Was the allocation sequence random?	PY	Quote: "those with PTSD Checklist–Civilian version (PCL-C) scores of at least 28 and a score of 3 or higher on 1 or more individual items were randomized, by a coin flip, into the exercise (MBX) group or the control (CON) group. Two researchers alternated flipping a coin, allowing the tossed coin to clatter to the floor. We chose to use this method for simplicity and convenience, although recent studies show that it may compromise the validity of
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	

			the randomization (27)"

"Another limitation may be treatment allocation bias in which the phlebotomy nurses could not be entirely blinded to the group assignment of participants" |
| | 1.3 Did baseline differences between intervention groups suggest a problem with the randomization process? | N | Quote: "There was no significant difference in age, ethnicity, education, marital status, smoking status, or nursing experience between the MBX and CON groups."

comment: PCL scores do not differ between groups [43.1 (11.2) vs. 42.6 (12.7)] |
	Risk of bias judgement	**Some concerns**	randomization was done by coin flipping; no information on allocation concealment; no baseline differences on relevant variables between groups
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	comment: instructor and participants were probably necessarily aware of the assigned condition as it is hardly possible to achieve blinding when psychological interventions are implemented
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	comment: no non-protocol interventions reported;
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	comment: no failures in implementation reported;

2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PN	Quote: "Participantscwere instructed to refrain from drinking alcohol, taking nonprescription drugs, and engaging in vigorous exercise for 72 hours before blood sampling. Compliance was verified by self-report."
		Quote: "As shown in Table 1, compliance among participants in the intervention was high, with 28 participants attending at least 75% of the 16 classes. More specifically, 1 participant (9%) attended 12 (75%) classes, 6 (55%) attended 13 (81%) classes, 3 (27%) attended 14 (88%) classes, and 1 (9%) attended 16 (100%) classes."
		comment: the amount of non-attendence described above is not likely to substantially affect participants' outcomes (i.e. lead to an underestimation of the treatment) effect
2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
Risk of bias judgement	**Low**	participants and intructor were probably aware of the assigned condition; high rate of adherence (28 out of 29 participants attended at least 75% of the classes); although a small risk of bias towards the null cannot be eliminated given the fact that some participants did not attend all classes, it is unlikely that participants' outcomes were substantially affected.

Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	Quote: "Twenty eight participants completed the study; 1 CON group member withdrew due to family problems (Figure 1)."
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	one participant in the control group withdrew due to family problems; data available for all other participants
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The PCL-C is a 17-item self-report instrument that measures the 17 Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) symptoms of PTSD, and it is commonly used to screen for PTSD (7). The instrument has been previously validated in individuals with PTSD, and it showed good test–retest reliability and internal consistency (30, 31)."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	comment: the PCL is a self-report questionnaire; thus, the 'assessors' were -in this case- the participants themselves (who were probably aware of their intervention)
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	Quote: "Based on over 20 years of experience in teaching martial arts for persons with high stress levels, the first author developed the intervention."
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	comment: there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered

			according to their beliefs/expectations about the intervention effect. It is particularly high as the comparator was a no-treatment condition (WL). Especially in light of the fact that the intervention was developed by the first author, lack of blinding is considered pariculary problematic. (no information regarding therapist allegiance)
	Risk of bias judgement	**High**	the risk of bias due to knowledge of the intervention is high, as the comparator is a passive control condition. In view of the fact that assessors were not blinded it should also be noted that the first author developped the intervention.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PY	Quote: "ClinicalTrials.gov Identifier: NCT01462045" comment: examination of the history of changes. a comparison of study record versions (clinicaltrials.gov) indicates that authors adhered to their pre-specified intentions in all aspects that are relevant for the result of interest
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	N	comment: information in the earliest study record version (before recruitment) is consistent with the reported measures, definitions, time points
	5.3 ... multiple eligible analyses of the data?	PN	comment: means and SDs for completers reported; generally, the risk of bias due selection of results based on multiple eligible analyses is lowered by the fact that results assessed here are raw values
	Risk of bias judgement	**Low**	comparison of study record versions (clinicaltrials.gov) indicates that authors adhered to their pre-specified intentions
Overall bias	**Risk of bias judgement**	**High**	unclear allocation concealment; no considerable deviations from intended interventions; data available for all participants but one in the control group; high risk of bias due to knowledge of the

					intervention; comparison of study record versions (clinicaltrials.gov) indicates that authors adhered to their pre-specified intentions with respect to the result of interest

Unique ID	85	Study ID	910101	Assessor	R
Ref or Label	Marks 1998	Aim	adhering to intervention (the 'per-protocol' effect)	The effect of adhering to intervention ...	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	CT	Comparator	REL	Source	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	PY	Quote: "Suitable patients gave written consent, were rated in a second 2-hour interview, and were then randomly assigned in permuted blocks 15 of 20 to undergo E, C, EC, or R, stratified for personal (intended by someone) or impersonal (eg, accidents) trauma. The therapist (K. L. or S. T.)then learned the patient's treatment condition."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PY	Quote: "The 4 treatment groups did not differ demographically. [...] The E group had the least current major depression (Table 1), and the E and R groups had lower baseline scores than the C group on 5 primary and 7 secondary measures and than EC on 7 primary and 7 secondary measures (1-way ANOVA with LSD paired comparisons)."

	Risk of bias judgement	**High**	no information on allocation concealment; significant baseline differences between groups on several clinical measures
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as it is hardly possible to achieve blinding when psychological interventions are implemented
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	N	Quote: "Treatments were delivered as planned, judged by blind independent ratings of their integrity."
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PN	Quote: "Of the 87 trial entrants, 74 (85%) consented to rating of audiotaped sessions by a "blind" behavioral-cognitive therapist outside the unit." "Treatment adherence was sound." "The therapist rated this [homework compliance] from patients' daily homework diaries, based on the percentage completion of the forthcoming week's homework negotiated at the end of each session. Mean (±SD) percentage compliance was as follows: all 77 patients, 63% ± 30%; [...] C, 43% ± 28%; [...], and R, 69%±28%. Lower compliance for C (x2=10.3, df=1, P=.02) could be artifactual, as challenging cognitions was harder to rate than time spent in E and R tasks."
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	no blinding of participants and therapists; no major deviations occurred; treatment fidelity and patient compliance were systematically assessed.

Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	Quote: "Of the 87 trial entrants, 10 (3 in the E group, 1 in the C group, 5 in the EC group, and 1 in the R group; P not significant) dropped out before becoming evaluable at week 6 (reasons seldom given)" comment: missing posttest data for n=1 out of n=19 randomized to CT, and for n=1 out of n=20 randomized to REL (5% in each group)
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	Data available for nearly all participants randomized (except for n=2 out of n=39).
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The PTSD measures were the Clinician-Administered PTSD Scale (CAPS 2) 21, 22 (assessor rated), which measured the frequency and intensity of 17 DSM-III-R PTSD symptoms [...]" comment: The administered scale (CAPS) is a validated PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "One assessor (a psychiatrist [H. N.] or a psychologist [M. L.]) screened each patient and rated him or her at weeks 0, 6, and 11 (posttreatment), and at 1-, 3- and 6-month follow-up thereafter. Assessors were kept unaware of the treatment condition."

			comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "Assessors were kept unaware of the treatment condition." comment: although the assessors were blind to the participants' condition, there was no blind assessment, since the participants (answering interview questions) might have been aware of their treatment allocation
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: assessment could have been influenced because participants were probably not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. This risk, however, is lowered by the fact that the comparator is also an active intervention.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN	
	Risk of bias judgement	**Some concerns**	participants were probably not blinded and might have answered according to their beliefs or expectations. This risk, however, is lowered by the fact that the comparator is also an active intervention.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PY	comment: three PTSD measures were used to assess symptom severity (CAPS, IES, PSS), but between-group results are only reported for the CAPS. Quote: "Two other PTSD measures were the

			Impact of Events Scale (IES) (self-rated), [...], and the PTSD Symptoms Scale (self), [...]"
	5.3 ... multiple eligible analyses of the data?	PN	the risk of bias due selection of results based on multiple eligible analyses is lowered by the fact that the results assessed here are raw values (means, SDs).
	Risk of bias judgement	**High**	no information regarding pre-specified analysis plan; results were not reported for all outcome measures of interest (not for IES and PSS); completers data reported (not for the ITT sample); generally, the risk of bias due selection of results based on multiple eligible analyses is lowered by the fact that the results assessed here are raw values (means, SDs). All in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are significant concerns.
Overall bias	**Risk of bias judgement**	**High**	no information on allocation concealment; significant baseline differences between groups on several clinical measures. no major deviations from intended interventions occurred. Data available for nearly all participants randomized (except for n=2). risk os bias due participants' (potential) knowledge of the intervention, however, lowered by the fact that the comparator is also an active intervention. no information regarding pre-specified analysis plan; results were not reported for all outcome measures of interest - therefore, the risk of bias is rated high.

Unique ID	88	**Study ID**	930101	**Assessor**	R

Ref or Label	Meffert 2014	Aim	adhering to intervention (the 'per-protocol' effect)	The effect of adhering to intervention ...	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	IPT	Comparator	WL	Source	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "Eligible participants were randomly assigned to IPT or waitlist control groups using a computer-generated random allocation sequence."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	NI	Quote: " The age of the IPT intervention group ranged from 21 to 42 years, with a mean of 31.3 years. The gender of the IPT intervention group was 83% women. The age of the waitlist control group ranged from 24 to 39 years, with a mean of 30.4 years. The gender of the waitlist group was 78% women." comment: no further information regarding baseline differences provided
	Risk of bias judgement	**Some concerns**	random allocation sequence; unclear allocation concealment; insufficient information regarding baseline differences.
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	Quote: "Participants were not blinded to group status. Given the cultural norms of frequent communication among Sudanese, it was not possible to prevent participants from being aware that they were receiving IPT at two different time points."
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	

			"It is important to note that neither the participants nor the therapists could be blinded to group status given the nature of the intervention."
2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?		PY	comment: no non-protocol interventions reported
2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?		PN	Quote: "Five members of the Sudanese community without prior mental health training were trained to deliver IPT." "Formal group supervision of IPT cases occurred twice per week, led by the first author. Informal supervision occurred nearly daily, through interactions related to screening of participants and administration of measures." comment: it should be noted that lay counsellors delivered therapy which might increase the probability of failures in implementation. However, no failures in implementing the intervention are reported.
2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?		PN	Quote: "Among the 22 randomized, 20 completed the protocol. There were no adverse events. One participant withdrew because her husband forbade her to continue. One dropped out secondary to time constraints." comment: one dropout per group. The proportion of participants who ceased treatment early is not considered high enough to cause significant bias;

			nearly all participants received the assigned intervention
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	unblinded participants and therapists; no failures in implementing the intervention are reported; high adherence (1 dropout per group).
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "Among the 22 randomized, 20 completed the protocol. There were no adverse events. One participant withdrew because her husband forbade her to continue. One dropped out secondary to time constraints." "1 Lost to follow-up [IPT group]" comment: amount of missing data: IPT n=2 [15.4%]; WL n=1 [11.1%]
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	Quote: " To address the effects of missing data, dropouts, and the one case lost to follow-up, we completed a last observation carried forward (LOCF) analysis" comment: no analysis correcting for bias or

			sensitivity analysis reported; the approach quoted above is not considered appropriate in order to correct for bias or to test robustness
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PN	Quote: "There were no adverse events. One participant withdrew because her husband forbade her to continue. One dropped out secondary to time constraints."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
			comment: equal proportion of missing data in both groups; reasons are probably unrelated to the outcome.
	Risk of bias judgement	**Low**	(small and) equal proportions of missing data in both groups; reasons are probably unrelated to the outcome.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The HTQ is a checklist developed by the Harvard Program in Refugee Trauma that has been used effectively with many refugee populations. [...] Cutoff scores for the HTQ have been developed to identify cases and noncases of PTSD. These cutoffs have been found to have greater than 90% sensitivity and specificity even when used without adaptation for culture differences (Ichikawa, Nakahara, & Wakai, 2006)."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points

	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	Quote: "Participants were not blinded to group status." "All measures were read to participants and their responses were recorded. The administrators of the measurements were the future (or former) therapists of the participants. Therapists were not blind to group status."
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	Quote: " It is important to note that neither the participants nor the therapists could be blinded to group status given the nature of the intervention. Related is the fact that the therapists administered measures to their own (future or past) patients. It is possible that both the participants and the therapists would have a bias toward appearing improved at the conclusion of IPT treatment or that both therapists and participants could have a bias toward reporting more symptoms prior to beginning IPT."
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	participants were aware of group status; the risk of bias due to knowledge of the intervention is particularly high because the comparator is a no-treatment condition; in addition, the interviewer knew the participants' treatment condition and was the participant's therapist - For subjective outcomes such as 'clinical impression of improvement', knowledge of the intervention received could be highly influential; risk of bias is very high.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	comment: no information on pre-specified analysis plan
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest

	5.3 ... multiple eligible analyses of the data?	PN	comment: results of interest are reported for all completers; generally, the risk of bias due to multiple eligible analyses of the data is low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure of interest; for all time points of interest; completers' results reported (not for the ITT sample); generally, the risk of bias due selection of results based on multiple eligible analyses is lowered by the fact that the results assessed here are raw values (means, SDs). All in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are some concerns.
Overall bias	**Risk of bias judgement**	**High**	random allocation sequence; unclear allocation concealment; insufficient information concerning baseline differences. unblinded participants and therapists; no failures in implementing the intervention are reported; high adherence (1 dropout per group). equal proportions of missing data in both groups; reasons are probably unrelated to the outcome. participants were aware of group status; the risk of bias due to knowledge of the intervention is particularly high because the comparator is a no-treatment condition; in addition, the interviewer knew the participants' treatment condition and was the participant's therapist - For subjective outcomes such as 'clinical impression of improvement', knowledge of the intervention received could be highly influential; risk of bias is very high. not enough information regarding pre-specified analysis plan to reliably assess the risk of bias due to selection of the reported result.

Unique ID	106	Study ID	1130102	Assessor	R
Ref or Label	Power 2002	Aim	adhering to intervention (the 'per-protocol' effect)	The effect of adhering to intervention ...	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	EMDR	Comparator	WL	Source	Journal article(s) with results of the trial
Outcome	self-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	NI	Quote: "Randomization was by means of a predetermined schedule unbeknown to the assessors, therapists or patients. Following completion of the entire initial assessment, for those patients who met entry criteria, the blind assessor then opened a sealed envelope that informed as to which group patients were to be allocated."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Y	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: " Similarly, as illustrated in Table 1 there were no differences between groups with regard to age, length of time since initial trauma, gender, marital status, history of previous psychiatric illness, type of trauma, or prescribed psychotropic medication at time of inclusion in the study."
	Risk of bias judgement	**Low**	incomplete description of the random sequence generation, without confirmation that there was a random component (see quote); allocation concealment; no baseline differences but no significance test reported for baseline PTSD scores (desciptively, they are higher in the EMDR group at baseline which might lead to a biased effect size estimate).
Bias due to deviations from	2.1 Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as blinding is not feasible when psychological interventions are implemented
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	

intended interventions	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	comment: no non-protocol interventions reported
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	NI	Quote: " Of the two therapists, one was more experienced in EMDR than the other." comment: no failures of implementation reported, but also no systematic assessment of therapist adherence/treatment fidelity
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	Y	Quote: "There was no formalized assessment of patient compliance with between session exposure homework requirements." "Drop-out rates between these three groups were as follows, 12 (31%) from EMDR, 16 (43%) from ECCR and five (17%) from WL"
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	PN	comment: naive 'per protocol' analysis used; no inverse probability weighting to adjust for censoring of participants who cease adherence to their assigned intervention
	Risk of bias judgement	**High**	no blinding of participants/therapists; no appropriate analysis used to estimate the effect of non-adherence (there were dropouts from treatment; patient compliance was not systematically assessed; neither was therapist adherence)
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: " Drop-out rates between these three groups were as follows, 12 (31%) from EMDR, 16 (43%) from EC-CR and five (17%) from WL (X^2 =5.6, df=2, p=0.06)."
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis reported
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: "D0.22. Comparisonbetweenthe33drop-outsandthe72 completers regarding presentation at time of initial assessment produced no significant
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	

| | | | differences on any of the demographic characteristics or treatment outcome measures with the sole exception of a higher frequency score on the CAPS-C Avoidance subscale for the drop-outs (t=2.2, df=103, p <0.05." "It may therefore be that those with high levels of avoidance of stimuli associated with the trauma are less likely to tolerate treatment approaches, whether EMDR or EC-CR,that entail some degree of confrontation with the traumatic image or situation" comment: proportions of dropouts: EMDR n=12 (31%), WL n=5 (17%), statistical significance of difference not tested; significance tests reported above (quote) include all three groups, not only those relevant here, however, results suggest that participants without posttest data had migher avoidance scores which incidactes that missingness might depend on its true value. reasons for dropout are not reported, making it difficult to assess the likeliness that missingness in the outcome depended on its true value. |
| | **Risk of bias judgement** | **High** | reasons for dropout not reported; descriptive difference in proportions of dropout (no significance test done/reported); participants without posttest data had higher avoidance scores; no sensitivity analysis or analysis correcting for bias. |

Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "A self-report version of the SI-PTSD Symptom Checklist (Davidson, Smith, & Kudler, 1989). This comprises 12 self-rated questions assessing the severity of DSM III-R symptoms each on a 0–4 scale. Three subscales can be derived from this measure relating to intrusive, avoidant and hyperarousal symptoms of PTSD." comment: The administered scale is a validated PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	participants were probably necessarily aware of the assigned condition as participant blinding is not feasible. The SI-PTSD-SR is a self-report questionnaire; therefore the 'assessors' (particiants) were probably aware of their intervention
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: assessment could have been influenced because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect comment: the risk of bias due to knowledge of the intervention is high, as the comparator is a passive control condition (here: WL).
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according

			to their beliefs/expectations about the intervention effect. It is particularly high as the comparator was a no-treatment condition (WL)
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	comment: no information provided
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	Quote: " CAPS assessments were not routinely collected on all control group subjects at end of WL period and therefore only the pre-treatment CAPS scores are presentedfor this group." comment: no justification provided why CAPS assessments were only done on some participants; this might raise some concerns. Apart from that point, results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: completers data reported; generally, the risk of bias due to multiple eligible analyses of the data is relatively low as the results assessed here are raw values (means, SDs)
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for all self-report outcome measures of interest (not for the CAPS at post-treatment, hence, slight concerns regarding selection of results based on results from multiple eligible outcome measures); for all time points of interest; results are reported for the completers sample; generally, the risk of bias due selection of results based on multiple eligible analyses is low as the results assessed here are raw values (means, SDs). All in all, difficult to assess due to lack of information; therefore, some concerns cannot be eliminated

Overall bias	Risk of bias judgement			High	insufficient information on random sequence generation; risk of bias due to participants' knowledge of intervention (measurement of outcome); potential bias due to missing outcome data as reasons for dropout are not reported and difference in avoidance symptoms between participants with/without missing posttest data; not enough information regarding pre-specified analysis plan to reliably assess the risk of bias due to selection of the reported result. no significance test reported for baseline PTSD scores (desciptively, they are higher in the EMDR group at baseline which might lead to a biased effect size estimate)

Unique ID	107	Study ID	1140101	Assessor	R
Ref or Label	Ratcliff 2016	Aim	adhering to intervention (the 'per-protocol' effect)	The effect of adhering to intervention …	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	MBI	Comparator	WL	Source	Journal article(s) with results of the trial
Outcome	self-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	NI	Quote: "Participants were randomly assigned to 1 of 3 groups: (1) yoga (YG); (2) stretching control (ST); or waitlist control (WL) using a form of adaptive randomization,46 according to age, stage of disease, time since diagnosis, type of surgery, and chemotherapy (neoadjuvant or adjuvant)."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	

	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "The 3 groups were similar on all medical and demographic variables (Table 1). There were no statistically significant differences among the groups on any of the self-reported variables at baseline, apart from the SF-36 general health subscale. Women in YG reported lower baseline general health compared with those in ST (P = .01)."
	Risk of bias judgement	**Some concerns**	
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	Quote: "Participants also were not blinded to study condition, [...]"
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	comment: no non-protocol interventions reported
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	NI	
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PY	Quote: " Out of a maximum possible 18 classes, 87% of YG and 85% of ST participants attended ≥12 classes (YG = 13.8; ST = 14.7). Only 3 patients in each group attended fewer than half the classes." "There were also no group, demographic, or baseline selfreport differences between those who completed the 6-month follow-up assessment and those who did not (Ps > .14)" comment: see CONSORT flow chart figure 1: assigned to MBI n=53, n=49 completed; assigned to WL n=54, n=48 completed waitlist period
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	PN	comment: naive 'per protocol' analysis used; no inverse probability weighting to adjust for censoring

			of participants who cease adherence to their assigned intervention
	Risk of bias judgement	**High**	neither participants nor instructors were blinded; participant non-adherence reported (n=10 dropouts) but no information regarding deviations from intended interventions reported beyond that; no appropriate analysis used to adjust for censoring of participants
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: " 13 dropped out before, and 15 after, randomization, for a baseline sample size of 163 (YG = 53; ST = 56; WL = 54)." comment: see CONSORT flow chart figure 1: assigned to MBI n=53, n=49 completed; assigned to WL n=54, n=48 completed post-assessments
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis reported
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: ""There were also no group, demographic, or baseline selfreport differences between those who completed the 6-month follow-up assessment and those who did not (Ps > .14), with the exception that older adults were more likely to complete the 6-month assessment."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	
			comment: equal proportions of dropout in both groups; no baseline self-report differences between participant with/without posttest data. However, reasons for dropout are not reported, so it is possible that reasons differed between groups and

			that reasons for dropout were study-related. Also, no information on the occurrence of adverse events is given
	Risk of bias judgement	**High**	7.5% missing posttest data in the Yoga group, 11.11% in the WL group; participants with and without posttest data did not differ on baseline self-report measures; however, no information on occurrence of adverse events and reasons for dropout are not reported which is why the risk of bias is rated high
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "Posttraumatic stress symptoms were measured by the Impact of Event Scale (IES), a scale that assesses the 2 most common categories of responses [...] Adequate internal reliability was found for the total scale (Cronbach's α = .85) as well as intrusive (Cronbach's α = .85) and avoidance (Cronbach's α = .79) subscales." comment: The administered scale (IES) is a validated PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	Quote: " Participants also were not blinded to study condition, and no measure of treatment expectation was collected, which could have biased the findings because of the subjective nature of the outcomes" comment: the IES is a self-report questionnaire;

			hence the 'assessors' were aware of their intervention
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	Quote: " Participants also were not blinded to study condition, and no measure of treatment expectation was collected, which could have biased the findings because of the subjective nature of the outcomes"
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
			comment: assessment could have been influenced because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect
			comment: the risk of bias due to knowledge of the intervention is high, as the comparator is a passive control condition (here: WL).
	Risk of bias judgement	**High**	there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. It is particularly high as the comparator was a no-treatment condition (WL)
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	comment: results were reported for the outcome measure of interest; for all time points of interest; however, based on the available information it is difficult to assess the risk of selective reporting, as it

				is impossible to know whether other (e.g. clinician-rated) PTSD measures were used but not reported
	5.3 ... multiple eligible analyses of the data?		NI	comment: only completers data reported (either no ITT analysis conducted or results not reported); generally, the risk of bias due to multiple eligible analyses of the data is lowered by the fact that the results assessed here are raw values (means, SDs)
	Risk of bias judgement		**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure of interest; for all time points of interest; completers' results reported (not for the ITT sample); All in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are some concerns.
Overall bias	**Risk of bias judgement**		**High**	limited information on randomization procedure and NI on allocation concealment; little information regarding possible deviations from the intended interventions (only dropout from treatments reported, NI on treatment fidelity, comliance, therapist effects or adverse events); risk of bias due to participants' knowledge of intervention; potential bias due to missing outcome data as reasons for dropout are not reported; not enough information regarding pre-specified analysis plan to reliably assess the risk of bias due to selection of the reported result.

Unique ID	108	**Study ID**	1170101	**Assessor**	R
Ref or Label	Rauch 2015	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention ...**	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants

Experimental	PE	Comparator	PCT	Source	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question		Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?		NI	Quote: "Veterans were randomly assigned to receive 10 to 12, 80-min sessions of PE or PCT."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?		NI	
				comment: The randomization procedure is described incompletely, without confirming that there was a random component. A simple statement such as "we randomly allocated" is considered insufficient to be confident that the allocation sequence was genuinely randomized. No information regarding allocation concealment
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?		NI	Quote: ".Table1 presents sample demographics."
				comment: Table 1 shows demographics for the total sample, for responders and for nonresponders but not on group-level, which raises concerns. No information regarding baseline differences reported beyond this table
	Risk of bias judgement		**Some concerns**	given that there is no information on random sequence generation, allocation concealment or baseline characteristics on group-level, there are serious concerns regarding the randomization process
Bias due to deviations from	2.1 Were participants aware of their assigned intervention during the trial?		PY	comment: therapists and participants were probably necessarily aware of the assigned condition as
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?		PY	

intended interventions			blinding is not feasible when psychological interventions are implemented
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NI	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	NI	Quote: " The first author served as the only study therapist." comment: the fact that the same therapist delivered all treatments may reduce the risk that there were major deviations but this is highly speculative as no information is reported that could be used to assess the risk of bias in this domain
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	NI	Quote: "Thirty-six veterans were consented (PE, n=18; PCT, n = 18; see CONSORT Flowchart and Checklist). Six of these veterans did not return for any study visit and no data is available. Twenty-six veterans completed treatment (PE, n=11, PCT, n=15; 87% retention)." comment: n=6 never started treatment (unclear in which group), n=4 ceased treatment early (unclear in which group). No further information regarding therapist or participant adherence
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NI	
	Risk of bias judgement	**High**	no information reported regarding deviations from the intended interventions beyond the fact that n=10 participants dropped out;

Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "Thirty-six veterans were consented (PE, n=18; PCT, n = 18; see CONSORT Flowchart and Checklist). Six of these veterans did not return for any study visit and no data is available. Twenty-six veterans completed treatment (PE, n=11, PCT, n=15; 87% retention)." comment: flowchart was not found in the report (or supplementary material), therefore, little information is available regarding missing data and dropouts
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis reported
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NI	comment: higher dropout rate in PE compared to the PCT group [PE: 38.89% ; PCT: 16.67% (statistical significance unclear)]; reasons for dropout unknown; amount of missing data is non-negligible. All in all, it is unclear whether missingness in the outcome depends on its true value, so there is a risk of bias
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	
	Risk of bias judgement	**High**	higher dropout rate in PE compared to the PCT group (statistical significance unclear); reasons for dropout are unknown; no information regarding differences between participants with and without missing data; amount of missing data is non-negligible. All in all, it is unclear whether missingness in the outcome depends on its true value.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "CAPS is a standard interview for PTSD severity. [...] The CAPS has excellent psychometrics."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points

	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "[pre-test] Evaluators were blind to veteran assignment" "[post-test] Independent evaluators completed CAPS interview and veterans completed self-report forms." comment: unclear whether interviewers conducting post-treatment assessments were blinded, too. participants were probably necessarily aware of the assigned condition, so either way there was no blinded assessment
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention. Thus, the probability that participants have answered according to their beliefs/expectations about the intervention effect is not very high (as opposed to a passive control condition)
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN	
	Risk of bias judgement	**Some concerns**	outcome assessors were probably aware of the intervention received; the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	
	5.3 ... multiple eligible analyses of the data?	NI	
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure of interest; for all time points of interest; completers' results reported (not for the ITT

			sample). All in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are some concerns.
Overall bias	**Risk of bias judgement**	**High**	no information on the randomization process or baseline differences between groups. insufficient information regarding deviations from the intended interventions. hardly any information on missing data. outcome assessors were probably aware of the intervention received; the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention. not enough information regarding pre-specified analysis plan to reliably assess the risk of bias due to selection of the reported result. Given the overall lack of information there are serious concerns that results might be biased.

Unique ID	109	**Study ID**	1180101	**Assessor**	R
Ref or Label	Reger 2016	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention …**	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	PE	**Comparator**	WL	**Source**	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	**Signalling question**	**Response**	**Comments**
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: " A research coordinator provided
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	their treatment group assignment based on computerized random number generation. Randomization was blocked in groups of three,

			such that one patient was assigned to each treatment group (PE, VRE, or WL) for every three participants enrolled."
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: " There were no statistically significant differences across the three groups in the distributions of demographic characteristics or in the baseline outcome measures."
	Risk of bias judgement	**Some concerns**	computerized random number generation, insufficient information on allocation concealment; no statistically significant differences across groups.
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as blinding is not feasible when psychological interventions are implemented
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	Quote: "Participants also had to agree not to initiate other psychotherapy for PTSD or new psychotropic medications during the treatment phase of the study." comment: no non-protocol interventions reported
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	Quote: "All therapy sessions were video recorded and 15% of planned sessions were randomly selected in advance for independent rating of treatment adherence and competence. Therapists were unaware of which sessions would be sent out for adherence review. Coders were not involved in other aspects of the study and were selected for this role based on experience as investigators on previous clinical trials of PE and VRE. Treatment adherence forms used in previous clinical trials of PE (Rothbaum, Astin, & Marsteller, 2005) were used for PE. [...] In 71 treatment sessions using

			prolonged exposure, 97.27% (962/989) of required criteria were observed."
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PY	Quote: "In 71 treatment sessions using prolonged exposure, 97.27% (962/989) of required criteria were observed." "By posttreatment, 44.44% of participants in the VRE group were lost to follow up or had withdrawn from the study compared to 40.74% of participants in the PE group [...]. Participants assigned to PE completed an average of 7.50 sessions (SD =3.46)." [out of ten sessions according to the intervention regimen]
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	N	comment: no inverse probability weighting to adjust for censoring of participants who cease adherence to their assigned intervention
	Risk of bias judgement	**High**	therapists and participants were probably necessarily aware of the assigned condition , therapist adherence systematically assessed and high; considerable amount of participant non-adherence (dropping out of treatment; non-attendance); no appropriate analysis to estimate the effect of adhering to the intervention
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "By posttreatment, 44.44% of participants in the VRE group were lost to follow up or had withdrawn from the study compared to 40.74% of participants in the PE group"

			comment: see figure 1 flow chart: in the PE group there is post-treatment data for n=32 out of n=54 participants randomized [40.74% missing data]; in the WL group for n=47 out of n=54 participants [12.96% missing data].
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	PN	Quote: "A key assumption of the linear mixed effects regression model is that the data were generated under a missing at random (MAR) or a covariate dependent assumption. Before estimating these models, we used a generalized linear model with a logit link and a Binomial error distribution to examine the association between the likelihood of dropout and several determinants, including CAPS scores, treatment assignment, and demographic variables. The results suggested that participants with lower education and those who did not identify as non-Hispanic White were more likely to drop out of the study during the treatment phase. Dropout was not related to CAPS scores. All regression models included education and race to improve the estimation. As a sensitivity analysis, we estimated a random coefficient selection model (Enders, 2010) that is appropriate for data that are missing not at random (MNAR). We specified a linear growth curve model for the first three measurement occasions using the CAPS last week." comment: the following aspects should be noted: (1) the results assessed here include only study completers (the quote refers to the ITT analysis), (2) given the very large amount of missing data, especially in the PE group, none of the approaches

			reported by the authors can fully eliminate the risk of bias.
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: ". Major reasons participants dropped out during treatment included geographic relocation away from the study site (WL [n= 4], PE [n =4], [...] time demands of military training/scheduling problems (WL [n= 0], PE [n= 1], [...] increases in symptomatology (WL [n =1], PE [n =1], [...] dissatisfaction with assigned treatment (WL [n =1], PE [n =4], [...] and losses to follow up (WL [n =1], PE [n= 7]"
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PN	
			comment: unequal proportions of dropout and differences in the reasons for dropout (e.g. more dissatisfaction with the treatment in the PE group; more losses to follow-up); overall is very considerable; the reported analyses indicate that results are not biased. Taken together, the reported

			analyses indicate that there are no differences between participants with and without posttest data that bias the result; however, these analyses do not address the potential relationship between missingness in the outcome and its true value in all aspects (e.g. differences in the reasons for dropout between the (PE/WL) groups), so it is possible that missingness in the outcome depends on its true value.
	Risk of bias judgement	**Some concerns**	very considerable amount of missing data; unequal proportions of dropout and differences in the reasons for dropout; sensitivity analysis conducted but not addressing all potential sources of bias; all in all, there are concerns that the result is biased due to missing outcome data.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The CAPS is a structured clinical interview that assesses the presence and severity of PTSD according to DSM–IV criteria." comment: The administered scale is a validated, gold-standard PTSD measure and likely to be sensitive to intervention effects

| 4.2 Could measurement or ascertainment of the outcome have differed between intervention groups? | PN | Quote: " All assessors were kept blind to patients' treatment group assignment through the use of assessment offices located in a separate hallway or a separate building relative to treating clinicians. Assessors were excluded from all study meetings involving discussions of clinical issues. Patients were instructed not to disclose their treatment group to the assessing clinicians. Assessors recorded accidental patient disclosures of treatment group and also guessed treatment group assignment at the end of each assessment. Patients broke the treatment group blind 29 times (WL =12, PE =5, VRE =12). Assessors guessed the correct treatment Group 53.8% of the time. At both the mid- and posttreatment assessments, over half of the correct guesses were WL, which likely reflects increased accuracy based on symptom presentation at the assessment. [...] . The intraclass correlation for CAPS severity was 0.94 at baseline using the last month reference period and 0.96 using the last week reference period. The intraclass correlation for CAPS severity at postassessment was 0.99 using the last week reference period. The intraclass correlation for PTSD diagnosis at baseline was 0.83 using CAPS last month reference period"

 "We should note that the waitlist period was 5-weeks though treatment typically took longer. We chose not to ask soldiers with PTSD to wait longer for treatment."

 comment: the same measurement methods and |

			thresholds were used for all participants. It should be noted, however, that time points probably differed to some extent between the groups (see quote)
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "Seven clinicians served as independent assessors, blind to participant treatment group assignment." "Patients broke the treatment group blind 29 times (WL =12, PE =5, VRE =12)." "At both the mid- and posttreatment assessments, over half of the correct guesses were WL, which likely reflects increased accuracy based on symptom presentation at the assessment. Restricting the guesses to only PE and VRE, the assessor was correct 36% of the time for the midpoint assessment, $x^2(2) =2.31$, $p =.31$ and 46% of the time for the posttreatment assessment, $x^2(2) = 5.11$, $p =.08$"

			comment: although the assessors were blind to the participants' condition, there was no blind assessment, since the participants (answering interview questions) were probably aware of their treatment allocation. In addition, the reported information indicates that blinding of interviewers was limited
4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY		comment: assessment could have been influenced because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect
4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY		comment: the risk of bias due to knowledge of the intervention is high, as the comparator is a passive control condition (here: WL). It would be lower if the comparator was another active intervention.
Risk of bias judgement	**High**		despite of the blinding of interviewers there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. The risk is particularly relevant as the comparator was a no-treatment waitlist condition.

| Bias in selection of the reported result | 5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis? | NI | Quote: " Therefore, we hypothesized that VRE and PE would reduce PTSD symptoms compared with a minimal attention waitlist. Furthermore, we hypothesized that VRE would significantly reduce PTSD symptoms compared with PE. We also hypothesized that soldiers assigned to VRE would demonstrate lower dropout rates, lower stigma and higher treatment satisfaction than soldiers assigned to PE."

"Funding for an additional recruitment site was received midway through the trial and the protocol was amended. In accordance with the documented plan for the grant at the end of the site's period of performance, the data from the original site were analyzed. This required a protocol deviation report to the institutional review board (IRB), as the amended protocol to add the second site was not sufficiently updated to reflect this planned analysis. However, once the findings were reviewed and presented to the IRB, the decision was made in collaboration with the IRB to halt recruitment at the second site and close the study. [...] .Only nine soldiers completed study participation from the second site. Accordingly, these soldiers were excluded from analyses and this article reports on all data collected at the primary study site."

"[...] contrary to our hypothesis, VRE was not superior to PE."

comment: transparent reporting of deviations from the intended study protocol and justifications - as |

			well as reporting that results did not support the authours' hypothesis- suggests that the risk of bias due to selection of the reported result is low. However, the only source used to assess this question is the journal article so the available information is very limited and may be insufficient to judge the risk of bias in this domain
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest (self- and clinician-rated PTSD); for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: both ITT and completers data reported;
	Risk of bias judgement	**Some concerns**	insufficient information regarding pre-specified analysis plan (but detailed reporting of study intentions, procedures, deviations from the intended protocols and the justifications offered may be regarded as indicative that the data was anlalysed in accordance with a prespecified analysis plan); results were reported for the outcome measure(s) of interest; for all time points of interest; results for both the ITT sample and completers reported;

			generally, the risk of bias due selection of results based on multiple eligible analyses is lowered by the fact that the results assessed here are raw values (means, SDs). All in all, it is difficult to assess the risk of bias in this domain due to lack of information.
Overall bias	**Risk of bias judgement**	**High**	computerized random number generation, insufficient information on allocation concealment; no statistically significant differences across groups. considerable amount of participant non-adherence (dropping out of treatment; non-attendance); no appropriate analysis to estimate the effect of adhering to the intervention. very considerable amount of missing data; unequal proportions of dropout and differences in the reasons for dropout; sensitivity analysis conducted but not addressing all potential sources of bias; all in all, there are concerns that the result is biased due to missing outcome data. risk of bias due to participants' knowledge of the intervention. not enough information regarding pre-specified analysis plan to reliably assess the risk of bias due to selection of the reported result.

Unique ID	111	**Study ID**	1190101	**Assessor**	R
Ref or Label	Reinhardt 2018	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention ...**	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	MBI	**Comparator**	WL	**Source**	Journal article(s) with results of the trial; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)

Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question			Response	Comments
Bias arising from the randomizatio n process	1.1 Was the allocation sequence random?			Y	Quote: "After participants completed baseline assessments, Author J. J. Noggle Taylor conducted a simple random assignment via a web-based random sampling service (www.randomizer.org; Urbaniak & Plous, 2015). Once we consented the minimum number of participants, we initiated baseline assessments and randomization to initiate a cohort."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?			NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?			NI	Quote: "Since attrition varied between groups, completer analysis of baseline differences in the primary outcome (CAPS) were evaluated using a two-way analysis of variance (ANOVA) of completer status, group assignment, and interaction between the two" "A two-way ANOVA comparing baseline differences in CAPS by completer status, group, and interaction showed no statistically significant differences in the overall model for any baseline CAPS outcomes (past-week lowest $p = 0.14$ [avoidance] and past-month lowest $p = 0.18$ [avoidance])." comment: based on the reported information (see quote) it is not clear whether all participants that were randomized are included in an analysis testing for baseline differences - including all relevant demographic and clinical variables.

	Risk of bias judgement	Some concerns	random assignment via a web-based random sampling service, no information regarding allocation concealment and insufficient information regarding baseline differences between groups (as initially randomized).
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	Quote: "Following randomization, participants were not blind to their group assignment."
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NI	Quote: "Ongoing medication or psychotherapy at enrollment was not exclusionary, and we requested notification of any changes to these regimens." comment: no further information regarding potential changes in medication during the trial or to what extent partiipants received, started or ceased other therapies outside of the study during the trial
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	NI	Quote: " all instructors (two female including author J. Johnston, and one male) had advanced training in Kripalu Yoga" comment: no information regarding failures in implementation or therapist adherence; it must be assumed that treatment fidelity was not systematically assessed

	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	Y	Quote: " During Stage 1, the dropout rate was 39% (20 of 51 participants randomized to start treatment). Over half of the yoga group (62%) [...] dropped out of the study. In contrast, dropout was less than one-quarter for the control group (16%)." "Participants were also asked to practice yoga outside of class daily for 15-minutes with a provided audio recording." comment: high attrition, especially in the yoga group (early cessation of treatment; non-attendence); it is not reported (or it was not assessed) to what extent participants complied with their homework to practice yoga
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	PN	comment: naive 'per protocol' analysis used; no inverse probability weighting to adjust for censoring of participants who cease adherence to their assigned intervention
	Risk of bias judgement	**High**	participants were not blinded; insufficient information regarding failures in implementation or instructor/participant non-adherence.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: " During Stage 1, the dropout rate was 39% (20 of 51 participants randomized to start treatment). Over half of the yoga group (62%) [...] dropped out of the study. In contrast, dropout was less than one-quarter for the control group (16%)." comment: the flow chart (figure 1) indicates that n= 21 participants in the control group completed stage 1 (10-week waiting period) and n=10 participants in the MBI group completed stage 1 (10-wk

			intervention). It is not reported why only data of n=6 WL participants [n=25 were randomized to control condition; amount of missing data: 76.0%] and n=9 yoga participants [n=26 were randomized to yoga; amount of missing data: 65.38%] is included in the completers analysis.
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis reported
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	Quote: "Thus, dropout was significantly lower in the control group compared to the yoga groups."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	PY	" This study's dropout is on the high end of ranges in published studies [...] Reasons for this may be because of the substantial time commitment for the yoga intervention, perceived heightened anxiety or triggers through enhanced mind–body awareness or interoception, and the use of a group yoga intervention format. Recruitment in this study was slow, and participant and yoga room scheduling conflicts were frequent. Future research might use individual sessions instead of group classes to possibly increase retention [...]" comment: the flow chart (figure 1) indicates that n= 21 participants in the control group completed stage 1 (10-week waiting period) and n=10 participants in the MBI group completed stage 1 (10-wk intervention). It is not reported why only data of n=6 WL participants [n=25 were randomized to control

			condition; amount of missing data: 76.0%] and n=9 yoga participants [n=26 were randomized to yoga; amount of missing data: 65.38%] were included in the analysis. proportions of dropout significantly differ between both groups. it is not reported whether participants with or without posttest data differed on any relevant variables. Reasons for dropout/missing data (or exclusion from the analysis?) are unknown. The amount of missing data is large. Taken together, the risk of bias due to missing outcome data is high.
	Risk of bias judgement	**High**	It is not reported why only data of n=6 WL participants [24%] and n=9 yoga participants [34%] were included in the analysis although there were more study completers; therefore it is unclear whether this data was missing or excluded for other reasons. unequal proportions of dropout between groups; no information on potential differences between participants with or without posttest data; insufficient information regarding reasons for dropout/missing data (or exclusion from the analysis?); overall amount of missing data is large; Taken together, the risk of bias due to missing outcome data is high.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The CAPS is a 30-item semistructured interview to diagnose PTSD. Subscales confirm criteria of (a) trauma and (b) PTSD symptoms of intrusion, (c) avoidance, and (d) hyperarousal." comment: The administered scale (CAPS) is a validated, gold-standard PTSD measure and likely to be sensitive to intervention effects

	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	Y	Quote: "Following randomization, participants were not blind to their group assignment." " For this study, two assessors (a female doctoral-level psychologist and a male psychiatry resident) conducted all CAPS interviews. They were both blinded to participant treatment condition." comment: the 'assessors' were -in this case- the participants themselves (who were aware of their intervention). According to the Cochrane guidelines, if either the participant is blinded and the data collector is not, or the data collector is blinded and the participant is not, then the outcome assessors should be considered to be aware of intervention received unless convincing evidence is available to the contrary.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	comment: although the interviewers were blinded, there is still a risk of bias. Because the comparator is a passive control condition the risk is higher that participants might have answered questions according to their beliefs/expectations regarding their assigned condition. This risk cannot be fully eliminated by assessment by a blinded clinician.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. It is particularly high as the comparator was a no-treatment condition (WL)

Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	PN	Quote: " We hypothesized that compared to a no-treatment assessment-only control group, PTSD symptoms would improve after a yoga intervention" "This study is registered at www.clinicaltrials.gov, #NCT00962403." comment: no SAP available; examination of the registry record reveals changes in the record that raise concerns - according to this information the trial started as a single-arm trial [which suggests that the reported hypothesis (quote) was changed post-hoc]: it was not until study completion that the number of arms was changed from "1" to "2" and allocation was changed from "non-randomized" to "randomized". In addition, the entry from August 2010 ['study status: recruiting'] does not include the PCL-C as a PTSD measure, whereas this instrument was added in March 2013 ['study status: completed']. In the final report (which is evaluated here) no mention of these substantial changes and no justification offered. There are concerns.
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	comment: results for the PCL-C are not reported. [authours state that results on this measure were identical which lowers the risk of bias.] As the result assessed here concerns only clinician-rated PTSD, the case mentioned is considered in the evaluation of results for self-rated PTSD [ID 112, 113].
	5.3 ... multiple eligible analyses of the data?	NI	comment: [see also 3.3] the flow chart indicates that n= 21 control group participants [out of n=25 randomized] and n=10 in yoga [n=26 randomized] completed. It is not reported why only data of n=6

			WL participants and n=9 yoga participants were included in the analysis.
	Risk of bias judgement	**Some concerns**	examination of the registry record (clinicaltrials.gov) reveals changes in the record which raise concerns; inconsistent information regarding the exclusion of participants from analysis raising additional questions; all in all, there are strong concerns about selective reporting.
Overall bias	**Risk of bias judgement**	**High**	random assignment via a web-based random sampling service, no information regarding allocation concealment and insufficient information regarding baseline differences between groups. participants were not blinded; insufficient information regarding deviations from intended interventions. Insufficient information on missing data and dropouts; large overall amount of missing data; available information indicates that risk of bias due to missing outcome data is high. risk of bias due to knowledge of the intervention because participants were not blinded and the comparator is a passive control condition. examination of the registry record (clinicaltrials.gov) reveals changes in the record which raise concerns; inconsistent information concerning the exclusion of participants from analysis raising additional questions; all in all, there are strong concerns with regard to selective reporting.

Unique ID	122	**Study ID**	1300101	**Assessor**	R

Ref or Label	Rothbaum 2005	Aim	adhering to intervention (the 'per-protocol' effect)	The effect of adhering to intervention ...	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	PE	Comparator	EMDR	Source	Journal article(s) with results of the trial; Personal communication with trialist
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomizatio n process	1.1 Was the allocation sequence random?	NI	Quote: "If the participant met criteria and gave consent, she was then randomized and scheduled accordingly: Seventy-four participants were randomly assigned to one of two active treatments (EMDR or PE) or a waitlist control group (WAIT)."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	NI	Quote: "In the completer sample of 60 women, mean participant age was 33.8 years (SD =11.0). [...] There were no significant differences among the three treatment conditions for any of these variables." ".Despite randomization,comparisons between the three groups at pretreatment revealed significant differences on some measures. As assessed by the CAPS, participants in the EMDR condition exhibited significantly higher overall PTSD symptoms [...] No differences between groups emerged on self-report measures of PTSD (PSS and IES-R) except that EMDR participants reported higher levels of intrusive symptoms on the PSS than did PE participants [...]. The EMDR group also exhibited significantly higher levels of depression [...] dissociation [...] and trait anxiety (STAI-T)" comment: baseline characteristics and pretreatment

			scores on clinical measures are only reported for completers (n=60) which rasies concerns. No information regarding baseline differences in the ITT sample (n=74)
	Risk of bias judgement	**Some concerns**	no information on random sequence generation or allocation concealment; baseline characteristics (demographic and clinical variables) only reported for completers sample, not for all participants randomized.
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	comment: therapists were probably necessarily aware and participants may have been aware of the assigned condition as blinding is not feasible when psychological interventions are implemented
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	comment: no non-protocol interventions reported

2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	Quote: "EMDR sessions were rated as 92.09% adherent for essential and unique items while PE sessions were rated 90.46% adherent for items considered essential to each protocol. Using a scale from 1 to 7, mean EMDR therapist skill was rated 6.04 (SD=0.58) or very good for essential and unique items. Mean PE therapist skill was rated 5.80 (SD=0.66) or very good for essential and unique items."
		comment: it should be noted that ratings were not done by independent researchers (Quote: "Dr. Edna Foa designated a PE expert from her lab to make these ratings for PE sessions") and researcher allegiance was not assessed; moreover, the investigators modified the instrument used to rate PE sessions without offering a justification or details regarding the modifications made; the scale used to rate EMDR was developed by the investigators and the EMDR rater themselves; this might raise concerns regarding the reliability
2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PY	Quote: "Of the 74 women enrolled in the study, 1 dropped out during the assessment phase, 1 was terminated and referred during treatment for not meeting treatment criteria, 12 dropped out during treatment, and 60 women (83.3%) completed the protocol"
2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	N	comment: naive 'per protocol' analysis used; no inverse probability weighting to adjust for censoring of participants who cease adherence to their assigned intervention

	Risk of bias judgement	High	treatment fidelity and therapist competence systematically assessed and rated as high, although there are concerns about the reliability of ratings; participant non-adherence (early dropouts) may have biased the result.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "Of the 74 women enrolled in the study, 1 dropped out during the assessment phase, 1 was terminated and referred during treatment for not meeting treatment criteria, 12 dropped out during treatment, and 60 women (83.3%) completed the protocol." "only 2 of 14 participants who did not complete the study (1 in each of the active treatments) provide data other than baseline" comment: no flow chart and limited information on participant flow
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	Quote: "The dropout rate across the three groups was not significantly different, PE: 13.0% (n =3,2 before MID); EMDR: 20.0% (n =5,4 before MID); and WAIT: 16.7% (n =4)."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	
			comment: equal proportions of dropout, no analysis addressing potential differences between participants with and without missing data; reasons for dropout are not reported; all in all, given the lack of information the risk of bias is high

	Risk of bias judgement	High	data of 16.7% of participants that were randomized missing; equal proportions of dropout, no analysis addressing potential differences between participants with and without missing data; reasons for dropout are not reported; in consideration of the limited information available the risk of bias is high.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: " For the current study, interrater reliability for the CAPS was 93.8% (κ = .79)." comment: The administered scale (CAPS) is a validated, gold-standard PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "Assessments were conducted at pretreatment, posttreatment, and follow-up of 6 and 12 months' posttreatment" comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "All assessments were conducted by IAs who were kept blind to the treatment condition." comment: participants may have been aware of the assigned condition so there was no blind assessment
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention. This significantly decreases the probability that
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN	

			participants answers differed based on differential beliefs/expectations about the intervention effect
	Risk of bias judgement	**Some concerns**	although the interviewers were blinded, assessment was not completely blind as participants were potentially aware of their allocation status; risk of bias due to knowledge of the intervention, however, is lowered by the fact that the comparator was also an active intervention.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	NI	comment: least square means are reported for completers, but no unadjusted results are available - and no details regarding the calculation of those LSMs are reported; authors offer a justification as to why no ITT analysis was conducted [Quote: "Because only 2 of 14 participants who did not complete the study (1 in each of the active treatments) provide data other than baseline, intent-to-treat analyses provide no consequentially different results and are not included here"].
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure of interest; for all time points of interest; completers' results reported but no anadjusted effect size estimates are reported and information on the calculation of least square means is limited; All in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are concerns.

Overall bias	Risk of bias judgement		High	no information on random sequence generation or allocation concealment; baseline characteristics (demographic and clinical variables) only reported for completers sample, not for all participants randomized. treatment fidelity and therapist competence systematically assessed and rated as high, although there are concerns about the reliability of ratings; participant non-adherence (early dropouts) may have biased the result. data of 16.7% of participants that were randomized missing; equal proportions of dropout, no analysis addressing potential differences between participants with and without missing data; reasons for dropout are not reported; in consideration of the limited information available the risk of bias is high. although the interviewers were blinded, assessment was not completely blind as participants were potentially aware of their allocation status; risk of bias due to knowledge of the intervention, however, is lowered by the fact that the comparator was also an active intervention. not enough information regarding pre-specified analysis plan to reliably assess the risk of bias due to selection of the reported result. participants in the EMDR condition exhibited significantly higher overall PTSD symptoms at baseline which might lead to a biased effect size estimate (SMD) for the CAPS

Unique ID	123	Study ID	1300102	Assessor	R
Ref or Label	Rothbaum 2005	Aim	adhering to intervention (the 'per-protocol' effect)	The effect of adhering to intervention ...	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants

Experimental	PE	Comparator	WL	Source	Journal article(s) with results of the trial; Personal communication with trialist
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question		Response	Comments
Bias arising from the randomizatio n process	1.1 Was the allocation sequence random?		NI	Quote: "If the participant met criteria and gave consent, she was then randomized and scheduled accordingly: Seventy-four participants were randomly assigned to one of two active treatments (EMDR or PE) or a waitlist control group (WAIT)."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?		NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?		NI	Quote: "In the completer sample of 60 women, mean participant age was 33.8 years (SD =11.0). [...] There were no significant differences among the three treatment conditions for any of these variables." ".Despiterandomization,comparisons between the three groups at pretreatment revealed significant differences on some measures. As assessed by the CAPS, participants in the EMDR condition exhibited significantly higher overall PTSD symptoms [...] No differences between groups emerged on self-report measures of PTSD (PSS and IES-R) except that EMDR participants reported higher levels of intrusive symptoms on the PSS than did PE participants [...]. The EMDR group also exhibited significantly higher levels of depression [...] dissociation [...] and trait anxiety (STAI-T)" comment: baseline characteristics and pretreatment scores on clinical measures are only reported for completers (n=60) which rasies concerns. No information regarding baseline differences in the ITT sample (n=74)

	Risk of bias judgement	Some concerns	no information on random sequence generation or allocation concealment; baseline characteristics (demographic and clinical variables) only reported for completers sample, not for all participants randomized.
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	comment: therapists were probably necessarily aware and participants may have been aware of the assigned condition as blinding is not feasible when psychological interventions are implemented
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	comment: no non-protocol interventions reported
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	Quote: "EMDR sessions were rated as 92.09% adherent for essential and unique items while PE sessions were rated 90.46% adherent for items considered essential to each protocol. Using a scale from 1 to 7, mean EMDR therapist skill was rated 6.04 (SD=0.58) or very good for essential and unique items. Mean PE therapist skill was rated 5.80 (SD=0.66) or very good for essential and unique items." comment: it should be noted that ratings were not done by independent researchers (Quote: "Dr. Edna Foa designated a PE expert from her lab to make these ratings for PE sessions") and researcher allegiance was not assessed; moreover, the investigators modified the instrument used to rate PE sessions without offering a justification or details regarding the modifications made; the scale used to rate EMDR was developed by the investigators and the EMDR rater themselves; this might raise concerns regarding the reliability

	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PY	Quote: "Of the 74 women enrolled in the study, 1 dropped out during the assessment phase, 1 was terminated and referred during treatment for not meeting treatment criteria, 12 dropped out during treatment, and 60 women (83.3%) completed the protocol"
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	N	comment: naive 'per protocol' analysis used; no inverse probability weighting to adjust for censoring of participants who cease adherence to their assigned intervention
	Risk of bias judgement	**High**	treatment fidelity and therapist competence systematically assessed and rated as high, although there are concerns about the reliability of ratings; participant non-adherence (early dropouts) may have biased the result.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "Of the 74 women enrolled in the study, 1 dropped out during the assessment phase, 1 was terminated and referred during treatment for not meeting treatment criteria, 12 dropped out during treatment, and 60 women (83.3%) completed the protocol." "only 2 of 14 participants who did not complete the study (1 in each of the active treatments) provide data other than baseline" comment: no flow chart and limited information on participant flow
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	

	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	Quote: "The dropout rate across the three groups was not significantly different, PE: 13.0% (n =3,2 before MID); EMDR: 20.0% (n =5,4 before MID); and WAIT: 16.7% (n =4)." comment: equal proportions of dropout, no analysis addressing potential differences between participants with and without missing data; reasons for dropout are not reported; all in all, given the lack of information the risk of bias is high
	Risk of bias judgement	**High**	data of 16.7% of participants that were randomized missing; equal proportions of dropout, no analysis addressing potential differences between participants with and without missing data; reasons for dropout are not reported; in consideration of the limited information available the risk of bias is high.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: " For the current study, interrater reliability for the CAPS was 93.8% (κ = .79)." comment: The administered scale (CAPS) is a validated, gold-standard PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "Assessments were conducted at pretreatment, posttreatment, and follow-up of 6 and 12 months' posttreatment" comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points

	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "All assessments were conducted by IAs who were kept blind to the treatment condition."
			comment: participants may have been aware of the assigned condition so there was no blind assessment
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	comment: there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. It is particularly high as the comparator was a no-treatment condition (WL)
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	although the interviewers were blinded, assessment was not completely blind as participants were potentially aware of their allocation status; risk of bias due to knowledge of the intervention is high because the comparator was a no-treatment WL condition.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	NI	comment: least square means are reported for completers, but no unadjusted results are available - and no details regarding the calculation of those LSMs are reported; authors offer a justification as to why no ITT analysis was conducted [Quote: "Because only 2 of 14 participants who did not complete the study (1 in each of the active treatments) provide data other than baseline, intent-to-treat analyses provide

			no consequentially different results and are not included here"].
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure of interest; for all time points of interest; completers' results reported but no anadjusted effect size estimates are reported and information on the calculation of least square means is limited; All in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are concerns.
Overall bias	**Risk of bias judgement**	**High**	no information on random sequence generation or allocation concealment; baseline characteristics (demographic and clinical variables) only reported for completers sample, not for all participants randomized. treatment fidelity and therapist competence systematically assessed and rated as high, although there are concerns about the reliability of ratings; participant non-adherence (early dropouts) may have biased the result. data of 16.7% of participants that were randomized missing; equal proportions of dropout, no analysis addressing potential differences between participants with and without missing data; reasons for dropout are not reported; in consideration of the limited information available the risk of bias is high. although the interviewers were blinded, assessment was not completely blind as participants were potentially aware of their allocation status; risk of bias due to knowledge of the intervention is high because the comparator was a no-treatment WL

condition. not enough information regarding pre-specified analysis plan to reliably assess the risk of bias due to selection of the reported result. participants in the EMDR condition exhibited significantly higher overall PTSD symptoms at baseline which might lead to a biased effect size estimate (SMD) for the CAPS

Unique ID	124	Study ID	1300103	Assessor	R
Ref or Label	Rothbaum 2005	Aim	adhering to intervention (the 'per-protocol' effect)	The effect of adhering to intervention ...	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	EMDR	Comparator	WL	Source	Journal article(s) with results of the trial; Personal communication with trialist
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	NI	Quote: "If the participant met criteria and gave consent, she was then randomized and scheduled accordingly: Seventy-four participants were randomly assigned to one of two active treatments (EMDR or PE) or a waitlist control group (WAIT)."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	

	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	NI	Quote: "In the completer sample of 60 women, mean participant age was 33.8 years (SD =11.0). [...] There were no significant differences among the three treatment conditions for any of these variables." ".Despiterandomization,comparisons between the three groups at pretreatment revealed significant differences on some measures. As assessed by the CAPS, participants in the EMDR condition exhibited significantly higher overall PTSD symptoms [...] No differences between groups emerged on self-report measures of PTSD (PSS and IES-R) except that EMDR participants reported higher levels of intrusive symptoms on the PSS than did PE participants [...]. The EMDR group also exhibited significantly higher levels of depression [...] dissociation [...] and trait anxiety (STAI-T)"
			comment: baseline characteristics and pretreatment scores on clinical measures are only reported for completers (n=60) which rasies concerns. No information regarding baseline differences in the ITT sample (n=74)
	Risk of bias judgement	**Some concerns**	no information on random sequence generation or allocation concealment; baseline characteristics (demographic and clinical variables) only reported for completers sample, not for all participants randomized.
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	comment: therapists were probably necessarily aware and participants may have been aware of the assigned condition as blinding is not feasible when psychological interventions are implemented
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	comment: no non-protocol interventions reported

2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	Quote: "EMDR sessions were rated as 92.09% adherent for essential and unique items while PE sessions were rated 90.46% adherent for items considered essential to each protocol. Using a scale from 1 to 7, mean EMDR therapist skill was rated 6.04 (SD=0.58) or very good for essential and unique items. Mean PE therapist skill was rated 5.80 (SD=0.66) or very good for essential and unique items."	
		comment: it should be noted that ratings were not done by independent researchers (Quote: "Dr. Edna Foa designated a PE expert from her lab to make these ratings for PE sessions") and researcher allegiance was not assessed; moreover, the investigators modified the instrument used to rate PE sessions without offering a justification or details regarding the modifications made; the scale used to rate EMDR was developed by the investigators and the EMDR rater themselves; this might raise concerns regarding the reliability	
2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PY	Quote: "Of the 74 women enrolled in the study, 1 dropped out during the assessment phase, 1 was terminated and referred during treatment for not meeting treatment criteria, 12 dropped out during treatment, and 60 women (83.3%) completed the protocol"	
2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	N	comment: naive 'per protocol' analysis used; no inverse probability weighting to adjust for censoring of participants who cease adherence to their assigned intervention	

	Risk of bias judgement	High	treatment fidelity and therapist competence systematically assessed and rated as high, although there are concerns about the reliability of ratings; participant non-adherence (early dropouts) may have biased the result.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	Quote: "Of the 74 women enrolled in the study, 1 dropped out during the assessment phase, 1 was terminated and referred during treatment for not meeting treatment criteria, 12 dropped out during treatment, and 60 women (83.3%) completed the protocol." "only 2 of 14 participants who did not complete the study (1 in each of the active treatments) provide data other than baseline" comment: no flow chart and limited information on participant flow
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Y	Quote: "The dropout rate across the three groups was not significantly different, PE: 13.0% (n =3,2 before MID); EMDR: 20.0% (n =5,4 before MID); and WAIT: 16.7% (n =4)."
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	
			comment: equal proportions of dropout, no analysis addressing potential differences between participants with and without missing data; reasons for dropout are not reported; all in all, given the lack of information the risk of bias is high

	Risk of bias judgement	High	data of 16.7% of participants that were randomized missing; equal proportions of dropout, no analysis addressing potential differences between participants with and without missing data; reasons for dropout are not reported; in consideration of the limited information available the risk of bias is high.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: " For the current study, interrater reliability for the CAPS was 93.8% (κ = .79)." comment: The administered scale (CAPS) is a validated, gold-standard PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "Assessments were conducted at pretreatment, posttreatment, and follow-up of 6 and 12 months' posttreatment" comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "All assessments were conducted by IAs who were kept blind to the treatment condition." comment: participants may have been aware of the assigned condition so there was no blind assessment
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y	comment: there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	

			intervention effect. It is particularly high as the comparator was a no-treatment condition (WL)
	Risk of bias judgement	**High**	although the interviewers were blinded, assessment was not completely blind as participants were potentially aware of their allocation status; risk of bias due to knowledge of the intervention is high because the comparator was a no-treatment WL condition.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	NI	comment: least square means are reported for completers, but no unadjusted results are available - and no details regarding the calculation of those LSMs are reported; authors offer a justification as to why no ITT analysis was conducted [Quote: "Because only 2 of 14 participants who did not complete the study (1 in each of the active treatments) provide data other than baseline, intent-to-treat analyses provide no consequentially different results and are not included here"].
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure of interest; for all time points of interest; completers' results reported but no anadjusted effect size estimates are reported and information on the calculation of least square means is limited; All in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are concerns.

Overall bias	Risk of bias judgement		High	no information on random sequence generation or allocation concealment; baseline characteristics (demographic and clinical variables) only reported for completers sample, not for all participants randomized. treatment fidelity and therapist competence systematically assessed and rated as high, although there are concerns about the reliability of ratings; participant non-adherence (early dropouts) may have biased the result. data of 16.7% of participants that were randomized missing; equal proportions of dropout, no analysis addressing potential differences between participants with and without missing data; reasons for dropout are not reported; in consideration of the limited information available the risk of bias is high. although the interviewers were blinded, assessment was not completely blind as participants were potentially aware of their allocation status; risk of bias due to knowledge of the intervention is high because the comparator was a no-treatment WL condition. not enough information regarding pre-specified analysis plan to reliably assess the risk of bias due to selection of the reported result. participants in the EMDR condition exhibited significantly higher overall PTSD symptoms at baseline which might lead to a biased effect size estimate (SMD) for the CAPS

Unique ID	128	Study ID	1310101	Assessor	R
Ref or Label	Rothbaum 1997	Aim	adhering to intervention (the 'per-protocol' effect)	The effect of adhering to intervention ...	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants

Experimental	EMDR	Comparator	WL	Source	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question		Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?		NI	Quote: "Twenty-one adult female victims of completed rape were randomly assigned to active treatment (EMDR) or a wait-list control group (WAIT)."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?		NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?		NI	Quote: " The demographic characteristics of the 18 completers are presented in Table 1." "There were no significant differences between groups on the demographic variables." comment: baseline characteristics only reported for completers sample, not for all participants randomized, and it is not reported whether there were group differences on any of the clinical variables.
	Risk of bias judgement		**Some concerns**	no information on random sequence generation or allocation concealment; baseline characteristics only reported for completers sample, not for all participants randomized, and it is not reported whether there were group differences on any of the clinical variables.
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?		PY	Quote: "At this evaluation, participants were evaluated by the IA as to inclusion and exclusion criteria, and the procedures of the study were explained in detail."
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?		PY	

			comment: therapists and participants were probably necessarily aware of the assigned condition as blinding is not feasible when psychological interventions are implemented, especially not in view of the apparent differences between an active and a passive treatment condition
2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?		PY	comment: no non-protocol interventions reported
2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?		PN	Quote: "Dr. Francine Shapiro designated one of her workshop leaders to review videotapes of treatment sessions to rate treatment integrity. [...] Integrityratingswere made ona 0-6 scale in which 0 was"unacceptable," 1 was "marginal," and 2-6 were "acceptable," [...] The 10 treatment integrity ratings averaged 3.9 (SD = 1.5), with a range from 2 to 6,indicating that the EMDR treatment delivered was deemed acceptable by an EMDR expert." comment: no substantial failures in implementation reported
2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?		PN	comment: see quote [question 2.4]: therapist adherence was systematically assessed; however, the Instrument appears to be rather vague as the scale does not assess specific Elements characteristic of the Intervention; the "expert rater" was involved in the study (Workshop leader) and researcher allegiance remains unknown; results should be interpreted with caution. In addition, three participants dropped out of treatment (14.28%)
2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?		NA	

	Risk of bias judgement	Some concerns	therapist adherence was systematically assessed and treatment fidelity was deemed acceptable by investigators; however, the Instrument used appears to be rather vague as the scale does not assess specific Elements characteristic of the Intervention; the "expert rater" was involved in the study (Workshop leader) and researcher allegiance remains unknown; In addition, three participants dropped out of treatment and no analysis adjusting for censoring of participants who cease adherence was conducted; although the number of dropouts is small, all in all, there are concerns.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	PN	Quote: "Eighteen participants completed the study; three participants completed only pretreatment assessment and then dropped out"
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: "Eighteen participants completed the study; three participants completed only pretreatmen tassessment and then dropped out: One found out she was pregnant immediately after assessment and decided not to continue; one assigned to WAIT decided to pursue private therapy and discontinued, and the third never attended her posttreatment assessment following WAIT"
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	
			comment: not reported to which group those 3 dropouts had been randomized (assumingly, at least 2 of them were assigned to WL), therefore proportions of dropout are unknown; however, the overall amount of missing data is relatively small (14.28%); pregancy is unrelated to the treatment condition, so this case does not lead to bias; hence,

			there are n=2 cases of missing data [9.52%] where missingness in the outcome could depend on its true value; it is unknown whether those 2 participants differed from completers on any demographic or clinical variables; in the light of the small amount of missing data, the risk of bias is not considered high but there are concerns.
	Risk of bias judgement	**Some concerns**	there are n=2 cases of missing data [9.52%] where missingness in the outcome could depend on its true value; one of those participants was assigned to the WL condition, group status of the other dropout is not reported; it is unknown whether those 2 participants differed from completers on any demographic or clinical variables; in the light of the small amount of missing data, the risk of bias is not considered high but there are concerns.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	PN	Quote: "The PTSD Symptom Scale (PSS; Foa, Riggs, Dancu, & Rothbaum, 1993) is a 17-item interview that corresponds to the 17 DSM-III-R and DSM-IV criteria for PTSD assessing the presence and severity of PTSD. [...] Cronbach's alpha calculated at the second administration on 118 subjects was .85. Interrater agreement was very good: kappa for diagnosis was .91." comment: The administered scale (PSS-I) is a validated PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "The posttreatment assessment was conducted after 4 weeks for all participants. [...] WAIT participants were assessed at the same 4-week time interval as EMDR subjects."

			comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "At this evaluation, participants were evaluated by the IA as to inclusion and exclusion criteria, and the procedures of the study were explained in detail." "Assessments were conductedpre-andposttreatment and 3 monthsfollowing treatment termination by an indepéndent assessor kept blind to treatment condition." comment: although the assessors were blind to the participants' condition, there was no blind assessment, since the participants (answering interview questions) were probably aware of their treatment allocation	
4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: assessment could have been influenced because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. the risk of bias due to knowledge of the intervention is particularly high as the comparator is a passive control condition.	
4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY		
Risk of bias judgement	**High**	although interviewers were blinded, there is a risk of bias due to knowledge of the intervention because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. It	

			is particularly high as the comparator was a no-treatment condition (WL).
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	NI	
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure of interest; for all time points of interest; completers' results reported (not for the ITT sample); generally, the risk of bias due selection of results based on multiple eligible analyses is lowered by the fact that the result assessed here are raw values (means, SDs). All in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are some concerns.

| Overall bias | Risk of bias judgement | | High | no information on random sequence generation or allocation concealment; baseline characteristics only reported for completers sample, not for all participants randomized, and it is not reported whether there were group differences on any of the clinical variables. treatment fidelity was deemed acceptable by investigators; however, there are doubts with regard to the reliability and validity of the Instrument used; researcher allegiance remains unknown; no analysis adjusting for censoring of participants who cease adherence (n=3) was conducted; although the number of dropouts was small, all in all, there are concerns. there are n=2 cases of missing data [9.52%] where missingness in the outcome could depend on its true value; one of those participants was assigned to the WL condition, group status of the other dropout is not reported; it is unknown whether those 2 participants differed from completers on any demographic or clinical variables; in the light of the small amount of missing data, the risk of bias is not considered high but there are concerns. interviewers blinded, but participants were not blinded and the risk of bias due to knowledge of the intervention is particularly high as the comparator was a no-treatment condition (WL). not enough information regarding pre-specified analysis plan to reliably assess the risk of bias due to selection of the reported result. pretreatment PTSD scores (PSS) for completers are higher in the WL group than in the EMDR group which leads to a biased effect size estimate (SMD(post-post)) |

Unique ID	139	Study ID	1430101	Assessor	R
Ref or Label	Shapiro 2015	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention ...**	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	EMDR	**Comparator**	WL	**Source**	Journal article(s) with results of the trial
Outcome	self-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	NI	Quote: "Participants were divided randomly from the list into two groups: eight participants in the first treatment group and nine in the waitlist/delayed treatment group."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	NI	Quote: "There was no statistically significant difference in age between the two groups (t[15] =0.92, p= .37)." "At the baseline assessment (T1), the mean IES-R scores for participants in immediate treatment (Group 1, M =41.63, SD =11.46) and waitlist/delayed treatment (Group 2, M =44.28, SD =17.53) were considered clinically significant for PTSD and were not different from each other (t[15] =0.36, p =.72). Initial mean PHQ-9 scores for Group 1 (M =13.13, SD =3.64) and Group 2 (M =10.11, SD =5.06) [...] were not significantly different from each other (t[15]=1.39, p =.18; see Table 1)." comment: no information on any other baseline characteristics beyond age and the two primary outcomes

	Risk of bias judgement	Some concerns	no information on the randomization process and on baseline differences beyond age and the two primary outcome variables.
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as blinding is not feasible when psychological interventions are implemented, especially not in view of the apparent differences between an active and a passive treatment condition
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PN	Quote: "Fidelity was not assessed, but all therapists were EMDR practitioners who had been trained in the R-TEP protocol." comment: therapist adherence was not systematically assessed and apart from the excerpt quoted no information is reported concerning deviations from the intended intervention. In view of the fact that only 2 sessions were administered and there were no dropouts, it is considered unlikely that there were other major deviations
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	Low	
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	

	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: " The IES-R commonly used clinically and in research is a self-report measure de signed to assess distress stemming from disturbing life events." comment: The administered scale (IES_R) is a validated PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: " The measurement scales were administered to participants by blind to protocol professionals by telephone. All participants were assessed at Time 1 (T1; see Figure 1). [...] One week later, participants from both groups were assessed again as before (Time 2 [T2])."
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	comment: participants were probably necessarily aware of the assigned condition as blinding is not feasible when psychological interventions are implemented, especially not in view of the apparent differences between an active and a passive treatment condition. Since the IES is a self-report measure, the (unblinded) participants were the assessors
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	Quote: "The sample comprised 17 neighbors and friends in the community, survivors of the fatal missile attack, who asked for psychological treatment after the incident." comment: assessment could have been influenced
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	

			because participants were not blind to their condition and might have answered according to their beliefs/expectations. The following aspects should be noted: (1) The fact that patients were personally aquainted with the PI -and actively initiated the treatment by asking him for help- makes it more likely that they were eager to please the investigators in return; the context of the study indicates that it was not difficult for participants to infer the purpose of the study, so there are serious concerns regarding demand effects; and (2) the risk of bias due to knowledge of the intervention is particularly high in view of the fact that the comparator was a passive control condition.
	Risk of bias judgement	**High**	unblinded assessors; two aspects suggest a very high risk of bias: (1) the comparator was a passive control condition; (2) patients were personally aquainted with the PI and actively initiated the treatment by asking him for help; because of lack of blinding there are serious concerns regarding demand effects.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	Quote: "The EMDR R-TEP (E. Shapiro & Laub, 2008, 2014; E. Shapiro, 2009, 2012) is a structured, comprehensive, and integrative recent trauma-focused protocol for EEI." "The sample comprised 17 neighbors and friends in the community, survivors of the fatal missile attack, who asked for psychological treatment after the incident." "In the wake of a sudden fatal missile attack on a town, this pilot project was organized to assist overwhelmed frontline mental health workers."

			"It was hypothesized that there would be significant differences between the treatment group and a waitlist group on the posttraumatic and depression scores" comment: the context of the trial suggests that (1) the primary author -who developed the intervention and who was aquainted with the patients- had an interest in finding EMDR to be superior to WL (researcher allegiance); and (2) that there may have been no time to plan the project well in advance, so it is unclear whether there was a pre-specified analysis plan.
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	
	5.3 ... multiple eligible analyses of the data?	PN	comment: Since (1) there was no missing data, (2) there were no baseline differences between both groups and (3) the methods used to calculate the raw values are straight forward, the risk that the numerical result was selected on the basis of results from multiple eligible analyses is regarded low
	Risk of bias judgement	**Some concerns**	information suggests that the study had not been planned long beforehand which raises the question whether there was a pre-specified analysis plan; researcher allegiance is likely; All in all, it is difficult to assess the risk of bias in this domain due to limited information; There are some concerns.

Overall bias	Risk of bias judgement		High	no information on the randomization process and on baseline differences beyond age and the two primary outcome variables. unblinded assessors; two aspects suggest a very HIGH RISK of bias: (1) the comparator was a passive control condition; (2) patients were personally aquainted with the PI and actively initiated the treatment by asking him for help; because of lack of blinding there are serious concerns regarding demand effects. information suggests that the study had not been planned long beforehand which raises the question whether there was a pre-specified analysis plan; researcher allegiance is likely; All in all, it is difficult to assess the risk of bias in this domain due to limited information; There are some concerns.

Unique ID	140	Study ID	1440101	Assessor	R
Ref or Label	Shapiro 2018	Aim	adhering to intervention (the 'per-protocol' effect)	The effect of adhering to intervention ...	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	EMDR	Comparator	WL	Source	Journal article(s) with results of the trial
Outcome	self-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	NI	Qoute: "after gaining informed consent to participate in the study, participants were divided randomly with names picked from a drum to construct the composition of the two groups with each person receiving a code number to conceal identities." "Whereas this study succeeded in randomizing the
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	

			allocation to groups, concealing selection and blinding outcome assessment and obtaining full data for the waiting list control parts of the study (T1 & T2), [...]" comment: although it is stated that allocation was random and "selection" was concealed, the information is considered insufficient to eliminate concerns; also, it is not confirmed that (1) it was not possible for people drawing names from the drum to see/identify the names of participants and (2) that cards with names were shuffled before drawing; similarly, no details are reported regarding allocation concealment. There are concerns
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "There were no statistically significant demographic or clinical differences in the two groups at baseline."
	Risk of bias judgement	**Some concerns**	insufficient information concerning random sequence generation and allocation concealment; no baseline differences.
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as blinding is not feasible when psychological interventions are implemented, especially not in view of the apparent differences between an active and a passive treatment condition
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PN	comment: no non-protocol interventions reported
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	NI	comment: no information on the therapist(s) delivering treatment, on treatment fidelity or regarding failures in implementation

	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	NI	Quote: "One female patient from group a dropped out of the study (age 41)." comment: no further information concerning therapist and paticipant adherence
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	PN	comment: no analysis estimating the effect of adhering to the intervention reported
	Risk of bias judgement	**High**	no information on the therapist(s) delivering treatment, on treatment fidelity, failures in implementation or participant adherence (except that one dropped out).
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	PY	Quote: " Of the thirteen participants who began treatment in group A, twelve were examined at T2 (92.3%)" "Of the twelve participants who began treatment B, all twelve were examined at T2 (100%)"
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The measures used included the PCL-5 measure of Post-traumatic symptoms derived from the Diagnostic and Statistical Manual of Mental Disorders, fifth Edition" comment: The administered scale (PCL) is a

			validated PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	comment: participants -the assessors (self-report)- were probably necessarily aware of the assigned condition as blinding is not feasible when psychological interventions are implemented, especially not in view of the apparent differences between an active and a passive treatment condition.
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: assessment could have been influenced because participants were not blind to their condition and might have answered according to their beliefs/expectations regarding the treatment efficacy or according to their beliefs about desired results (to please the investigators; demand effects). the risk of bias due to knowledge of the intervention is particularly high since the comparator was a passive control condition.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	self-report; unblinded participants; the risk of bias due to knowledge of the intervention is particularly high because the comparator was a no-treatment control condition.

Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	Quote: "The EMDR R-TEP, first published in 2008 (Shapiro & Laub 2008, 2014), is a structured, comprehensive, and integrative recent trauma-focused protocol for Early EMDR Intervention (EEI)."
			" The sample comprised twenty-five residents of the town, exposed to the intensive rocket attacks, who asked for psychological treatment after the two-month long flare-up of hostilities."
			"It was hypothesized: 1- there would be significant reduction in the posttraumatic and depression measures between the treated intervention group A and the waitlist control group B [...]"
			comment: the context of the trial suggests that the primary author -who developed the intervention- had an interest in finding EMDR to be superior to WL (concerns regarding researcher allegiance); no information on the pre-specified analysis plan.
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	
	5.3 ... multiple eligible analyses of the data?	NI	
	Risk of bias judgement	**Some concerns**	researcher allegiance is considered to be likely; no information on pre-specified analysis plan.
Overall bias	**Risk of bias judgement**	**High**	insufficient information concerning random sequence generation and allocation concealment; no baseline differences. no information on the therapist(s) delivering treatment, on treatment fidelity, failures in implementation or participant adherence (except that one dropped out). self-report; unblinded participants; the risk of bias due to

knowledge of the intervention is particularly high because the comparator was a no-treatment control condition. researcher allegiance is considered to be likely; ; not enough information regarding pre-specified analysis plan to reliably assess the risk of bias due to selection of the reported result.

Unique ID	143	Study ID	1520101	Assessor	R
Ref or Label	Thorp 2019	Aim	adhering to intervention (the 'per-protocol' effect)	The effect of adhering to intervention ...	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	PE	Comparator	Relaxation	Source	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	NI	Quote: " Male combat veterans (N = 87; mean age = 65 years) were
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	randomly assigned to 12 sessions of PE (n = 41) or RT (n = 46)."
			comment: The randomization procedure is described incompletely, without confirming that there was a random component. A simple statement such as "we randomly allocated" is

			considered insufficient to be confident that the allocation sequence was genuinely randomized.
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	Quote: "Demographic comparisons between treatment groups are located in Table 1. Veterans who received PE had served longer in the military than veterans who received RT (t(57) = 2.53, p < .05). There were no differences on pre-treatment symptoms between the treatment groups, including the CAPS (t(85) = -2.19, p = .77, d = .47), PCL-S (t(80) = -1.00, p = .29, d = .22), or PHQ-9 (t(79) = -2.28, p = .12, d = .40), nor were there differences in pre-treatment demographics."
	Risk of bias judgement	**Some concerns**	insufficient information regarding random sequence generation; no information on allocation concealment; no substantial baseline differences on any demographic or clinical variables.
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as blinding is not feasible when psychological interventions are implemented
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	NI	Quote: "Individuals taking psychotropic medications were allowed to participate, but those who had made changes to type or dosage level within 60 days were asked to delay their start until their medication regimen had stabilized for two months." comment: participants' medication status was assessed at baseline; but no information concerning non-protocol interventions during the study; e.g. assessment of changes in medication status during the study, seeking another therapy outside of the study (assumingly, it was not systematically assessed)

	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	Quote: "Therapists were adherent to both protocols, as measured by average percentage of required elements successfully administered during each session (PE = 91.73%; RT = 88.00%; t(24) = 2.41, p = .20)."
			comment: no failures in implementation reported
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PY	Quote: " Therapists were adherent to both protocols, as measured by average percentage of required elements successfully administered during each session (PE = 91.73%; RT = 88.00%; t(24) = 2.41, p = .20)."

"Treatment completion rates did not significantly differ between the two

conditions; 73% of individuals in the PE condition completed all 12 sessions of treatment (M = 9.39, SD = 4.51), while 78% of those who received RT completed all 12 sessions (M = 9.89, SD = 4.18; t(85) = -.49, p = .38)." |
| | 2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention? | N | comment: no inverse probability weighting to adjust for censoring of participants who cease adherence to their assigned intervention |
| | **Risk of bias judgement** | **High** | unblinded ('open') study; high therapist adherence but significant amount of participants who ceased treatment early; no analysis to adjust for censoring of participants who did not adhere. |
| **Bias due to missing outcome data** | 3.1 Were data for this outcome available for all, or nearly all, participants randomized? | N | Quote: "Study attrition rates were 31% at post-treatment and 48% at

follow-up; rates did not differ between treatment groups (p > .05)." |

	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	Quote: "Study attrition rates were 31% at post-treatment and 48% at
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	follow-up; rates did not differ between treatment groups (p > .05)."
			"Allocated to PE intervention (n = 41) [...] Completed first post-treatment assessment (e.g., CAPS, SCID; n = 29); Lost to follow-up (n = 12)"
			"Allocated to RT intervention (n = 46) [...] Completed first post-treatment assessment (e.g., CAPS, SCID; n = 37); Lost to follow-up (n = 9)"
			comment: amount of missing data =29.26% in PE, =19.56% in Relaxation; it is not reported whether participants with and without posttest data differed on any study variables; reasons for dropout are not reported; therefore, there is a risk of bias.
	Risk of bias judgement	High	
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: " The CAPS addresses each of the 17 symptoms from the Diagnostic and Statistical Manual for Mental Disorders – 4th Edition Text Revision (DSM-IV-TR; American Psychiatric Association, 2000), and it has high levels of internal consistency, good inter-rater reliability, and excellent convergent validity (Weathers, Keane, & Davidson, 2001)."

	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "Participants completed the PCL-S and PHQ-9 at the beginning of every therapy visit. Therapists scored each questionnaire in session and discussed changes from prior sessions and overall trajectories with the participant."
			comment: the fact that participants' answers were discussed with the unblinded therapist in every session may have influenced participants' responses in consecutive assessments (see 4.5); however, the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: " The ICEs [independent clinical evaluators] were masked to condition, were trained to administer the assessments by a doctoral-level psychologist and masters-level clinician, and received weekly supervision/consultation."
			comment: although the assessors were blind to the participants' condition, there was no blind assessment, since the participants were not blinded
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the outcome is not based solely on self-ratings of unblinded participants and that the comparator was also an active intervention. This significantly decreases the probability that participants' answers differed based on differential beliefs/expectations about the intervention effect. However, it should be noted that unblinded therapists discussed participants'
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	

			answers on the PCL in every session which may have influenced their answers in consecutive assessments; this raises strong concerns.
	Risk of bias judgement	**High**	The risk of bias due to lack of participant blinding is lowered by the fact that (1) it was not reled solely on self-ratings of PTSD symptom severity and (2) that the comparator was also an active intervention; interpretation of participants' self-reported answers on the PCL questionnaire and discussion of those answers with an unblinded therapist in every treatment session may have influenced subsequent measurements.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	NI	comment: the result of interest (means, SDs) is not reported for the ITT sample; raw values for completers are reported.
	Risk of bias judgement	**Some concerns**	no information regarding pre-specified analysis plan; results were reported for the outcome measure of interest; for all time points of interest; completers' results reported (not for the ITT sample); All in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are some concerns.

Overall bias	Risk of bias judgement		High	insufficient information regarding random sequence generation; no information on allocation concealment; no substantial baseline differences on any demographic or clinical variables. unblinded ('open') study; high therapist adherence but significant amount of participants who ceased treatment early; no analysis to adjust for censoring of participants who did not adhere. The risk of bias due to lack of participant blinding is lowered by the fact that (1) it was not reled solely on self-ratings of PTSD symptom severity and (2) that the comparator was also an active intervention; however, interpretation of participants' self-reported answers on the PCL questionnaire and discussion of those answers with an unblinded therapist in every treatment session may have influenced subsequent measurements. not enough information regarding pre-specified analysis plan to reliably assess the risk of bias due to selection of the reported result.

Unique ID	147	Study ID	1580101	Assessor	R
Ref or Label	Vaughan 1994	Aim	adhering to intervention (the 'per-protocol' effect)	The effect of adhering to intervention ...	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	EMDR	Comparator	WL	Source	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1
Domain	**Signalling question**			**Response**	**Comments**

Bias arising from the randomization process	1.1 Was the allocation sequence random?	NI	Quote: "After assessment each subject was randomly assigned to a treatment group and also to a wait list or nonwait list group. The procedure resulted in unequal numbers of subjects in the treatment groups - 12 in EMD, 13 in IHT and 11 in AMR. Seventeen of the 36 were initially assigned to the wait list and were reassessed after 2-3 weeks (mean 18.2 days, SD 0.8) before undergoing their active treatment." comment: unclear whether assignment to WL was random; generally, it is not confirmed that there was a random component.
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "Treatment groups did not differ on number of treatment sessions, demographic, trauma-related or symptom variables except that GAD was over-represented in IHT (77%) compared to AMR (27%) (x2, p = .035)."
	Risk of bias judgement	**Some concerns**	insufficient information regarding random sequence generation; no information on allocation concealment; no substantial baseline differences on any demographic or clinical variables.
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as blinding is not feasible when psychological interventions are implemented, especially not in view of the apparent differences between an active and a passive treatment condition
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	comment: no non-protocol interventions reported
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	NI	comment: no information regarding treatment fidelity/therapist cometence and adherence;

	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	NI	comment: no information on therapist or participant adherence; either adherence (and generally, deviations from the intended intervention) not systematically assessed or not reported
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	PN	comment: 'per protocol' analysis used; no inverse probability weighting to adjust for censoring of participants who cease adherence to their assigned intervention
	Risk of bias judgement	**High**	therapists and participants unblinded; no information regarding deviations from the intended intervention
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	PY	comment: no dropouts and no missing data is reported (nor explicitly reported that there was no missing data); information in table 2 indicates that all participants randomized completed posttests
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	no dropouts and no missing data is reported (nor explicitly reported that there was no missing data); information in table 2 indicates that all participants randomized completed posttest assessment.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	PN	Quote: "The PTSD Structured Interview (SI-PTSD; Davidson, Smith & Kudler. 1989), which scores each of 17 DSM-III-R criteria for severity on a scale of O-4, was administered." comment: The administered scale (SI-PTSD) is a validated PTSD measure and likely to be sensitive to intervention effects

	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "Structured interview. This was conducted by a blind rater (RG) on four occasions during the study - initial entry, wait list reassessment (17 individuals), posttreatment and follow-up." comment: although the assessors were blind to the participants' condition, there was no blind assessment, since the participants (answering interview questions) were probably aware of their treatment allocation
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: assessment could have been influenced because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. the risk of bias due to knowledge of the intervention is high, as the comparator is a passive control condition (WL)
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	NI	
	Risk of bias judgement	**High**	the risk of bias due to knowledge of the intervention is high as participants were not blinded and the comparator was a no-treatment control condition.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	comment: no information regarding pre-specified analysis plan; as the paper was published (year 1994) before guidelines for reporting study procedures and results as well as pre-registrations of trials were promoted, little information is available to assess the risk of bias due to selective reporting
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	comment: results are not reported for the IES for the waitlist group, only for the EMDR group
	5.3 ... multiple eligible analyses of the data?	NI	

	Risk of bias judgement		Some concerns	no information regarding pre-specified analysis plan; since the paper was published (year 1994) before clinical guidelines for reporting study procedures and results became popular (and before pre-registration of trials was promoted), little information is available to assess the risk of bias due to selective reporting.
Overall bias	**Risk of bias judgement**		High	insufficient information regarding random sequence generation; no information on allocation concealment; no substantial baseline differences on any demographic or clinical variables. therapists and participants unblinded; no information regarding deviations from the intended intervention. the risk of bias due to knowledge of the intervention is high as participants were not blinded and the comparator was a no-treatment control condition. in view of the publication year (1994), little information is available to assess the risk of bias due to selective reporting.

Unique ID	148	**Study ID**	1600101	**Assessor**	R
Ref or Label	Wahbeh 2016	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention ...**	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	MBI	**Comparator**	Relaxation	**Source**	Journal article(s) with results of the trial
Outcome	self-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the	1.1 Was the allocation sequence random?	PY	Quote: "Allocation was determined with a covariate adaptive randomization approach to ensure arms were well matched on important baseline characteristics and to reduce selection bias (Cai,
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	

randomization process			Xia, Xu, Gao,& Yan, 2006; Pocock & Simon, 1975)."
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	NI	Quote: "A total of 102 combat veterans completed the study (Figure1). Participants were mostly college educated male combat veterans from the Vietnam Era. There were no significant differences on important demographic variables (Table1)." "No Significant Differences Between Arms on Important Baseline Characteristics" [including PTSD scores] comment: reported only for completers
	Risk of bias judgement	**Some concerns**	no information on allocation concealment; baseline characteristics (demographic and clinical variables) only reported for completers, not for all participants randomized.
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	Y	Quote: "Participants were informed of their assignment by the research assistant (RA) at the first training visit." "Participants were trained one on one by an unblinded and trained RA once a week for 6 weeks."
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	comment: no non-protocol interventions reported; efforts were made to ensure that medication status, amount of excercise and other activities that could affect psychological or physiological measures remained stable

2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	Quote: "Each intervention was structured to maintain equipoise and reduce performance bias." "instructor evaluations were the same between arms (all ps > 0.05). Participants positively endorsed the instructor being confident about the intervention, comfortable working with them and enthusiastic about the intervention (MM − 4.78 ± 0.05, SB − 4.80 ± 0.13, MM+SB−4.49±0.05, SQ−4.88±0.13; p > 0.05; 5=strongly agree.)" comment: the reported information does not indicate that there were failures in implementation; it should be noted, though, that therapist adherence was not systematically assessed
2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PY	Quote: "Most participants were adherent. Average adherence levels ranged from 15 to 30 minutes per day, with the SQ arm participants practicing more than 20 minutes per day." "MM was different than SB on BDI (p=0.03), PSQI (p=0.009), GIC (p < 0.00005), and adherence (p= < 0.00005)" comment: 12 dropouts; it is not reported to which group they had been assigned
2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	PN	comment: naive 'per protocol' analysis used; no inverse probability weighting to adjust for censoring of participants who cease adherence to their assigned intervention

	Risk of bias judgement	**High**	participants and instructor unblinded; no formal assessment of treatment fidelity; participant adherence systematically assessed and acceptable, but 12 participants dropped out during the study (unclear how many of them were in the two intervention groups assessed here); No appropriate analysis was used to estimate the effect of adhering to the intervention.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	N	comment: see figure 1: n=12 dropouts who are not included in the analysis, their group allocation status is not reported, so the amount of missing data for the two intervention groups assessed here is unknown. If equal number of dropouts per group, then data missing for 10.34%
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	N	comment: no analysis correcting for bias or sensitivity analysis
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	PY	comment: beyond the number of noncompleters there is no information; it is unknown if (1) proportions of missing data are equal across groups; (2) participants with and without missing data differ on any study variables; (3) reasons for dropout were related to the treatment. Therefore, there is a high risk of bias.
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NI	
	Risk of bias judgement	**High**	beyond the number of noncompleters there is no information; it is unknown if (1) proportions of missing data are equal across groups; (2) participants with and without missing data differ on any study variables; (3) reasons for dropout were related to the treatment. Therefore, there is a high risk of bias.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The following measures were used as pre-post evaluative measures: PTSD Checklist (PCL; Weathers, Litz, Herman, Huska, &Keane, 1993); [...]"

			comment: The administered scale (PCL) is a validated PTSD measure and likely to be sensitive to intervention effects
4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN		Quote: "All participants had a telephone screening, screening visit, baseline visit, six training visits, and an endpoint visit. All visits occurred in the Oregon Health & Science University Hatfield Research Center."
4.3 Were outcome assessors aware of the intervention received by study participants?	Y		Quote: "Participants were informed of their assignment by the research assistant (RA) at the first training visit." comment: the PCL is a self-report questionnaire; hence the assessors were aware of their group allocation
4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Y		comment: assessment could have been influenced because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. this risk is lowered by the fact that the comparator was also an active intervention.
4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN		
Risk of bias judgement	**Some concerns**		measurement could be biased because participants were not blind to their condition; this risk is lowered by the fact that the comparator was also an active intervention.

Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	Quote: "This study's primary goal was to explore MM's mechanism of action in combat veterans with PTSD by separately examining two common components of structured MM programs, slowed breathing and mindfulness concepts. We proposed three pathways by which these components may potentially improve clinical outcomes [...]" comment: no pre-registration of the trial; no SAP available to the RoB assessor; little information regarding pre-specified analysis plan
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	
	5.3 ... multiple eligible analyses of the data?	NI	Quote: "In general, the four arms were assessed for differences with and without adjustment for covariates." "A completer rather than intention-to-treat analysis was conducted because of the mechanistic study aims." comment: raw values (means, SDs) (=unadjusted effect size estimates) are reported for completers; justification for analytic strategy (PP instead of ITT analysis) is appropriate.
	Risk of bias judgement	**Some concerns**	no prospective registration of the trial; insufficient information regarding pre-specified analysis plan to assess the risk of bias; reported justification for analytic strategy (PP instead of ITT analysis) seems appropriate; All in all, it is difficult to assess the risk of bias in this domain due to lack of information; There are some concerns.

Overall bias	Risk of bias judgement		High	no information on allocation concealment; baseline characteristics (demographic and clinical variables) only reported for completers, not for all participants randomized. participants and instructor unblinded; treatment fidelity not systematically assessed; participant adherence of completers was high, but 12 participants dropped out during the study (unclear how many of them were in the two intervention groups assessed here); no appropriate analysis used to estimate the effect of adhering to the intervention. beyond the number of noncompleters there is no information; it is unknown if (1) proportions of missing data are equal across groups; (2) participants with and without missing data differ on any study variables; (3) reasons for dropout were related to the treatment. Therefore, there is a high risk of bias. measurement could be biased because participants were not blind to their condition; this risk is lowered by the fact that the comparator was also an active intervention. not enough information regarding pre-specified analysis plan to reliably assess the risk of bias due to selection of the reported result.

Unique ID	150	Study ID	1620102	Assessor	R
Ref or Label	Wells 2015	Aim	adhering to intervention (the 'per-protocol' effect)	The effect of adhering to intervention ...	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	PE	Comparator	WL	Source	Journal article(s) with results of the trial; Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
Outcome	self-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomizatio n process	1.1 Was the allocation sequence random?	Y	Quote: "Stratified permuted block randomization was used giving equal allocation to each group and stratified by patient gender. Randomization lists were generated by the statistician prior to commencing the study and the therapists were blind to the randomization lists which were held and administered by individuals who were independent of the study."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	Quote: "Table 1 summarizes diagnoses, number of traumas, duration of PTSD, and demographic data." " No significant differences were found for chronicity of PTSD symptoms or BDI scores (but BDI scores approached significance: p = .06). Other pre-treatment differences on outcome measures were tested with a series of one-way ANOVA's which revealed no differences."
	Risk of bias judgement	**Some concerns**	Stratified permuted block randomization; insufficient information on allocation concealment; no substantial baseline differences across groups.
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	Quote: "Participants were aware upon consenting to take part that they may be randomly allocated to a condition that involved an 8-week waiting period with no intervention during this time." "A significant limitation is that blinding procedures were not possible and the therapists acted as assessors which may have introduced bias."
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Y	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	Quote: " Participants who had been prescribed psychotropic medication were required to maintain a stable medication regime throughout the treatment period."

	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	Quote: "Treatment adherence ratings were calculated as a percentage indicating the number of treatment components identified as a proportion of the total number of components possible if perfect adherence was achieved. The two therapists achieved 93 and 92 % adherence for PE"
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PN	Quote: "Treatment adherence ratings were calculated as a percentage indicating the number of treatment components identified as a proportion of the total number of components possible if perfect adherence was achieved. The two therapists achieved 93 and 92 % adherence for PE" " Completers attended 8 sessions of therapy or an 8-week waiting period. Of the 30 completers two individuals attended 7 rather than 8 MCT sessions and then completed the end of treatment measures." comment: high therapist adherence, all PE completers had attended all sessions (one participant was referred at mid treatment)
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	high therapist adherence (93 and 92%); excellent participant adherence.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	Quote: " Thirty participants completed the trial, one Exposure participant was referred on after the mid-point assessment" "Allocated to Exposure condition: 7 males 4 females (n = 11)

			Analyzed end treatment Completers (n = 10)" "Allocated to WL condition: 6 males 4 females (n = 10) Analyzed end wait Completers (n = 10)"
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	missing posttest data for n=1 participant.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "The Post-traumatic Stress Diagnostic Scale (PDS; Foa 1995) was the secondary outcome measure, used in addition to the IES because it assesses DSM-IV symptom dimensions. Test–retest reliability has shown good agreement (kappa .74) with high reliability across two time periods (87.3 %). Cronbach alpha of .92 was calculated for items 22–38 (corresponding to DSM-IV criteria b, c and d) indicating that the Symptom Severity Score calculated from these items has internal consistency." comment: The administered scale (PDS) is a validated PTSD measure and likely to be sensitive to intervention effects

	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "All self-report measures were administered at pre-treatment, post-treatment and at 3-month follow-up. Heart-rate was assessed at pre-treatment and post treatment. The IES was also administered at mid-treatment (end of session 4) to assess possible differences in rate of change."
			comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "Participants were aware upon consenting to take part that they may be randomly allocated to a condition that involved an 8-week waiting period with no intervention during this time."
			""A significant limitation is that blinding procedures were not possible"
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: assessment could have been influenced because participants were not blind to their condition and might have answered according to their beliefs/expectations about the intervention effect. the risk of bias due to knowledge of the intervention is high, as the comparator is a passive control condition (WL).
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	measurement could be biased because participants were not blinded; the risk of bias due to knowledge of the intervention is particularly high in view of the fact that the comparator was a passive control condition.

Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	Quote: "International Clinical Trials Registry Platform NHSTCT Register: ISRCTN63706856." comment: the trial was registered (http://www.isrctn.com/ISRCTN63706856), however, the registry record indicates that registration was done retrospectively (record: "Prospective/Retrospective: Retrospectively registered") so there is limited information on a pre-specified analysis plan
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	
	5.3 ... multiple eligible analyses of the data?	PN	Quote: " 2). These analyses were followed by pair-wise between-group (Bonferroni) comparisons on the adjusted means. Descriptive (unadjusted) statistics for the pre-treatment, post-treatment and follow-up data are presented in Table 2." comment: unadjusted and adjusted effect size estimates for completers are reported; raw values (means, SDs)
	Risk of bias judgement	**Some concerns**	the trial was registered at the ISRCTN registry; however, registration was done retrospectively; insufficient information on a pre-specified analysis plan to reliably assess the risk of bias in this domain.
Overall bias	**Risk of bias judgement**	**High**	insufficient information on allocation concealment. measurement could be biased because participants were not blinded; the risk of bias due to knowledge of the intervention is particularly high in view of the fact that the comparator was a passive control condition. the trial was registered at the ISRCTN registry; however, registration was done retrospectively; insufficient information on a pre-

specified analysis plan to reliably assess the risk of bias in this domain.

Unique ID	151	Study ID	1670101	Assessor	R
Ref or Label	Zang 2014	Aim	adhering to intervention (the 'per-protocol' effect)	The effect of adhering to intervention ...	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	NET	Comparator	WL	Source	Journal article(s) with results of the trial
Outcome	self-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "Thirty participants were randomly allocated to either NET (n = 10), NET-R (n = 10) or a waiting list condition (WL; n = 10) by a computer-generated list of random numbers."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: "There were no significant differences among the three groups regarding age, gender, education, marital status, income, injury, and house damage. Table 3 shows the mean scores of scales, except for brief COPE, of three groups at each time point (T1, T2, T3, and T4). At baseline (T1), there was no significant difference among three groups."
	Risk of bias judgement	**Some concerns**	computer-generated list of random numbers; no information on allocation concealment; no baseline differences on relevant variables between groups
	2.1 Were participants aware of their assigned intervention during the trial?	PY	

	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	Quote: "All participants gave informed consent after receiving a full explanation of the study design, objectives and explicit information regarding what the study entailed." comment: therapists and participants were probably necessarily aware of the assigned condition as blinding is not feasible when psychological interventions are implemented, especially not in view of the apparent differences between an active and a passive treatment condition
Bias due to deviations from intended interventions	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	Quote: "No major deviation from the study protocol occurred."
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	N	Quote: "Treatment adherence was monitored by the direct observation of treatment sessions, by case discussions in supervision meetings, and by a review of the records and treatment protocols." "All participants completed the treatment. No major deviation from the study protocol occurred."
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	Quote: "Treatment adherence was monitored by the direct observation of treatment sessions, by case discussions in supervision meetings, and by a review of the records and treatment protocols." "There were no drop-outs, with all participants completing the entire course of treatment and follow-up." " All Participants of NET condition spent no more than one session on narrating previous traumatic events, and then used two to three sessions

			focused on the single incident of the earthquake. They completed the treatment with four sessions in two weeks."
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	no blinding; no deviations from the study protocol and excellent participant adherence.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	Quote: "There were no drop-outs, with all participants completing the entire course of treatment and follow-up."
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	no missing data.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "Severity of PTSD symptoms was assessed using the Impact of Event Scale-Revised (IES-R; [29]). [...] Cronbach alpha for the three subscales of the Chinese IES-R have been reported as between .83-.89"
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points

	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "All participants gave informed consent after receiving a full explanation of the study design, objectives and explicit information regarding what the study entailed." comment: participants were probably necessarily aware of the assigned condition as blinding is not feasible when psychological interventions are implemented, especially not in view of the apparent differences between an active and a passive treatment condition
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: assessment could have been influenced because participants were not blind to their condition (e.g. demand effects; beliefs/expectations regarding treatment effects).the risk of bias due to knowledge of the intervention is high, as the comparator is a passive control condition.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	self-report; no blinding; the risk of bias due to knowledge of the intervention is high as participants were not blinded and the comparator was a no-treatment control condition.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	comment: no information of interest was missed in the report; detailed report of study procedures; deviations; data analysis and measurement procedures; therefore the overall probability of selective reporting is considered low; however, no pre-specified analysis plan is available so it is not possible to reliably assess the risk of bias in this domain
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest

Domain	Signalling question	Response	Comments
	5.3 ... multiple eligible analyses of the data?	PN	comment: results reported for all participants that had been enrolled; Since (1) there was no missing data, (2) there were no baseline differences on the measure between groups and (3) the methods used to calculate the raw values are straight forward, the risk that the numerical result was selected on the basis of results from multiple eligible analyses is regarded low
	Risk of bias judgement	**Some concerns**	detailed report of study procedures; deviations; data analysis and measurement procedures; therefore, selective reporting is not considered to be likely; however, no pre-specified analysis plan is available so it is not possible to reliably assess the risk of bias in this domain.
Overall bias	**Risk of bias judgement**	**High**	no information on allocation concealment. self-report; no blinding; the risk of bias due to knowledge of the intervention is high as participants were not blinded and the comparator was a no-treatment control condition. based on the available information selective reporting is not likely but there is not enough information regarding the pre-specified analysis plan to reliably assess the risk of bias in this domain.

Unique ID	152	**Study ID**	1680101	**Assessor**	R
Ref or Label	Zang 2013	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention ...**	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	NET	**Comparator**	WL	**Source**	Journal article(s) with results of the trial
Outcome	self-rated PTSD	**Results**	SMD(between)	**Weight**	1
Domain	**Signalling question**			**Response**	**Comments**

Bias arising from the randomization process	1.1 Was the allocation sequence random?	Y	Quote: "Twenty two participants were randomly allocated to either NET (n=11) or a waiting list condition (WL; n=11) by a computer-generated list of random numbers."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	N	Quote: " There were no significant differences between the two groups regarding age, gender, education, marital status, income, injuries, and house damage." "Table 2 showed the mean scale scores of two groups at each time point (T1, T2, T3, and T4). At baseline (T1), there is no significant difference between two groups."
	Risk of bias judgement	**Some concerns**	computer-generated list of random numbers; no information on allocation concealment; no baseline differences.
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	Quote: "All participants gave informed consent after receiving a full explanation of the study design and objectives and explicit information regarding what the study entailed."
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	comment: therapists and participants were probably necessarily aware of the assigned condition as blinding is not feasible when psychological interventions are implemented, especially not in view of the apparent differences between an active and a passive treatment condition
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	Quote: " No major deviation from the study protocol was apparent."
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	N	Quote: "Treatment adherence was monitored by the direct observation of treatment sessions, by case discussions in supervision meetings, and by a

			review of the records and treatment protocols." "All participants completed the treatment. No major deviation from the study protocol was apparent."
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	N	Quote: "Treatment adherence was monitored by the direct observation of treatment sessions, by case discussions in supervision meetings, and by a review of the records and treatment protocols." "All participants completed the treatment. No major deviation from the study protocol was apparent." "There were no drop-out, with all participants completing the entire course of treatment and follow-up."
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	no blinding; no deviations from the study protocol and excellent participant adherence.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	Quote: "There were no drop-out, with all participants completing the entire course of treatment and follow-up."
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	no missing data.
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	N	Quote: "Severity of PTSD symptoms was assessed using the Impact of Event Scale-Revised (IES-R; [26])."

			comment: The administered scale (IES-R) is a validated PTSD measure and likely to be sensitive to intervention effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "All participants gave informed consent after receiving a full explanation of the study design and objectives and explicit information regarding what the study entailed." comment: participants were probably necessarily aware of the assigned condition as blinding is not feasible when psychological interventions are implemented, especially not in view of the apparent differences between an active and a passive treatment condition
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: assessment could have been influenced because participants were not blind to their condition (e.g. demand effects; beliefs/expectations regarding treatment effects).the risk of bias due to knowledge of the intervention is high, as the comparator is a passive control condition.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	self-report; no blinding; the risk of bias due to knowledge of the intervention is high as participants were not blinded and the comparator was a no-treatment control condition.
Bias in selection of	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	comment: no information on the pre-specified analysis plan is available so it is not possible to reliably assess the risk of bias in this domain

the reported result	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	comment: results were reported for all outcome measures of interest; for all time points of interest
	5.3 ... multiple eligible analyses of the data?	PN	comment: results reported for all participants that had been enrolled; Since (1) there was no missing data, (2) there were no baseline differences on the measure between groups and (3) the methods used to calculate the raw values are straight forward, the risk that the numerical result was selected on the basis of results from multiple eligible analyses is regarded low
	Risk of bias judgement	**Some concerns**	no information regarding a pre-specified analysis plan; the risk that the numerical result was selected on the basis of the results from multiple eligible analyses of the data is relatively low in the context of the trial.
Overall bias	**Risk of bias judgement**	**High**	no information on allocation concealment. self-report; no blinding; the risk of bias due to knowledge of the intervention is high as participants were not blinded and the comparator was a no-treatment control condition. not enough information regarding the pre-specified analysis plan to reliably assess the risk of bias in this domain.

Unique ID	154	**Study ID**	40101	**Assessor**	R
Ref or Label	Alghamdi 2015	**Aim**	adhering to intervention (the 'per-protocol' effect)	**The effect of adhering to intervention ...**	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	NET	**Comparator**	WL	**Source**	Journal article(s) with results of the trial
Outcome	self-rated PTSD	**Results**	SMD(between)	**Weight**	1

Domain	Signalling question	Response	Comments
Bias arising from the randomizatio n process	1.1 Was the allocation sequence random?	Y	Quote: "Thirty four traumatized fire-fighters were randomly allocated to either NET (n ¼ 17) or a waiting list condition (WLC; n ¼ 17) by a computer generated list of random numbers."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	Quote: " There were no significant differences between the two groups regarding age, level of education, traumatic events number, or marital status, but the years of service were significantly higher in the WLC group. There were also no significant differences between the two groups (NET & WLC) in the mean scale scores of PTSD, anxiety, depression, coping strategies, and social support at the baseline test."
	Risk of bias judgement	**Some concerns**	computer generated list of random numbers; no information on allocation concealment; no substantial baseline differences between groups.
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	Quote: "The purpose of this study was described to the participants by the first author, and the procedure for collecting data was explained."

comment: therapists and participants were probably necessarily aware of the assigned condition as blinding is not feasible when psychological interventions are implemented, especially not in view of the apparent differences between an active and a passive treatment condition |
| | 2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial? | PY | |
| | 2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups? | PY | comment: no non-protocol interventions reported |

	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	Quote: " Participants spent no more than one session on narrating previous traumatic events, with 2e3 sessions focused on the index events available associated with the firefighters work met the DSM-IV criteria E for PTSD. All participants completed the treatment. No major deviation from the study protocol was apparent."
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PN	Quote: " Participants spent no more than one session on narrating previous traumatic events, with 2e3 sessions focused on the index events available associated with the firefighters work met the DSM-IV criteria E for PTSD. All participants completed the treatment. No major deviation from the study protocol was apparent."
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	no blinding; no major deviations from intended interventions.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	Quote: " There was no attrition across T1, T2, and T3 [...]" "All participants completed the treatment."
	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	no missing data.

Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	PN	Quote: "The prevalence rate of PTSD symptoms was assessed by using the Scale of Posttraumatic Stress Symptoms (SPTSS) (Carlson, 2001). This scale has been used widely for assessing PTSD symptoms in three subscales (re-experience, hyper arousal and avoidance), and the Arabic version was validated by Jaber (2012). [...] For the current study sample, Cronbach's alpha scores indicated acceptable internal consistency, being 90, .80, .78, and .73 for the total scale, re-experience, avoidance, and hyper-arousal subscales respectively."
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	Quote: "Fig. 1 presents the research and treatment schedules for both conditions." "The post intervention assessment was conducted on 20 January 2013 for the NETgroup and on 18 February 2013 for the WLC group." comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: "The purpose of this study was described to the participants by the first author, and the procedure for collecting data was explained." comment: participants (self-report, therefore they are the assessors) were probably necessarily aware of their group status and even of the study purpose which may influence their answers

	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	comment: participants were unblinded and even of the study purpose which may have influenced their answers (beliefs regarding treatment efficacy; demand effects); the risk of bias due to participant's knowledge of the intervention is particularly high because the comparator was a passive control condition.
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PY	
	Risk of bias judgement	**High**	participants were informed about the study purpose and study procedures; the risk of bias due to participant's knowledge of the intervention is particularly high because the comparator was a passive control condition.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	comment: no information beyond the reported study hypotheses
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	NI	
	5.3 ... multiple eligible analyses of the data?	PN	comment: Since (1) there was no missing data, (2) there were no baseline differences on the PTSD measure between groups and (3) the methods used to calculate the raw values are straight forward, the risk that the numerical result was selected on the basis of results from multiple eligible analyses is considered low
	Risk of bias judgement	**Some concerns**	insufficient information regarding pre-specified intentions/analysis plan; it is unlikely that the numerical result was selected on the basis of the result from multiple eligible analyses.
Overall bias	**Risk of bias judgement**	**High**	no information on allocation concealment. participants were informed about the study purpose and study procedures; the risk of bias due to participant's knowledge of the intervention is particularly high because the comparator was a passive control condition. not enough information regarding pre-specified analysis plan to reliably

					assess the risk of bias due to selection of the reported result.

Unique ID	155	Study ID	170102	Assessor	R
Ref or Label	Carlson 1998	Aim	adhering to intervention (the 'per-protocol' effect)	The effect of adhering to intervention ...	occurance of non-protocol interventions; failures in implementing the intervention that could have affected the outcome; non-adherence to their assigned intervention by trial participants
Experimental	EMDR	Comparator	REL	Source	Journal article(s) with results of the trial
Outcome	clinician-rated PTSD	Results	SMD(between)	Weight	1

Domain	Signalling question	Response	Comments
Bias arising from the randomization process	1.1 Was the allocation sequence random?	NI	Quote: " Following pretreatment assessment, the participants were assigned randomly to one of three conditions."
	1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	NI	
			comment: no further details provided
	1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	PN	Quote: " no significant differences between the groups with re- spect to ethnic makeup, marital status, employment (employed/not em- ployed), or recent psychiatric hospitalization. The groups did differ on age, largely owing to one veteran in the EMD group who was over 70 years old at the time of treatment, $F(2, 32) = 5.05$, $p < .01$. (This veteran was included owing to his Vietnam combat experience and the limited size of this group.)"

	Risk of bias judgement	Some concerns	no information on random sequence generation or allocation concealment; no substantial baseline differences.
Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial?	PY	comment: therapists (and participants) were necessarily aware of the assigned condition
	2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	PY	
	2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?	PY	Quote: " a 12-session treatment regimen was followed, otherwise adhering strictly to the EMDR treatment protocol recommended by Shapiro (1995)." " One author in addition to the therapist was in attendance during all treatment sessions, in the role of technician only, to monitor treatment and to main- tain a computerized count of numbers of eye movement periods. In addi- tion, the therapists used the standard EMDR manual session checklist to structure the EMDR treatment. These data were reviewed on a session- by-session basis. Repeated team consultation regarding each case and re- view of selected cases assured adherence to the treatment protocol." comment: no non-protocol interventions reported; adherence to protocol stated
	2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?	PN	Quote: " a 12-session treatment regimen was followed, otherwise adhering strictly to the EMDR treatment protocol recommended by Shapiro (1995)." In addi- tion, the therapists used the standard EMDR manual session checklist to structure the

			EMDR treatment. These data were reviewed on a session- by-session basis. Repeated team consultation regarding each case and re- view of selected cases assured adherence to the treatment protocol." comment: adherence and successful implementation reported by authors
	2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	PN	Quote: "Participants in the RXT group were also given a recorded cassette containing modified progressive relaxation instructions, written instructions for daily home use, and a self-monitoring form for re- cording usage. (In light of a very low return rate of the self-monitoring form, these data were not analyzed.)" comment: apart from low return rate of self-monitoring forms for PMR practice, no non-adherence reported. It is not considered likely that this factor leads to significant risk of bias
	2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	NA	
	Risk of bias judgement	**Low**	no blinding; high therapist adherence reported; no dropouts from treatment.
Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomized?	Y	Quote: "After group assignment and prior to posttreatment assessment, no participants dropped out of the control or eye movement groups and one participant dropped out of the biofeedback-assisted relaxation group (an attrition rate of 3% for all the assigned groups)."

	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	NA	
	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	NA	
	3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	NA	
	Risk of bias judgement	**Low**	no missing data
Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	PN	Quote: "In several 1- to 2-hr sessions, psychometric instruments were admin- istered including the following: The Clinician Administered PTSD Scale, CAPS-1 (Blake et al., 1995); [...]" comment: The CAPS is a validated, gold-standard PTSD measure which is likely to be sensitive to treatment effects
	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?	PN	comment: the same measurement methods and thresholds were used for all participants, and used at comparable time points
	4.3 Were outcome assessors aware of the intervention received by study participants?	PY	Quote: " To enhance objectivity of assessment, the psychometric instruments were normally administered by one of two research assistants blind to treatment group assignment." comment: the statement that interviews were 'normally' conducted by blind assessors raises concerns that this was not always the case. Moreover, assessments were not blinded either way, as participants might have been aware of their group status
	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	PY	

|			
	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	PN	comment: the risk of bias due to knowledge of the intervention is lowered by the fact that the comparator was also an active intervention. Thus, the probability that participants answered according to their beliefs/expectations about the intervention effect is not very high (as opposed to a passive control condition)
	Risk of bias judgement	**Some concerns**	blind interviewers, unblinded participants, risk of bias due to knowledge of the intervention is lowererd by the fact that the comparator was also an active intervention.
Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	NI	
	5.2 ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	PN	
	5.3 ... multiple eligible analyses of the data?	PN	comment: raw values (means, SDs) (=unadjusted effect size estimates) are reported for all completers.
	Risk of bias judgement	**Some concerns**	
Overall bias	**Risk of bias judgement**	**Some concerns**	blind interviewers, unblinded participants, risk of bias due to knowledge of the intervention is lowererd by the fact that the comparator was also an active intervention.

Appendix P

Comparison-adjusted funnel plot for the initial NMA
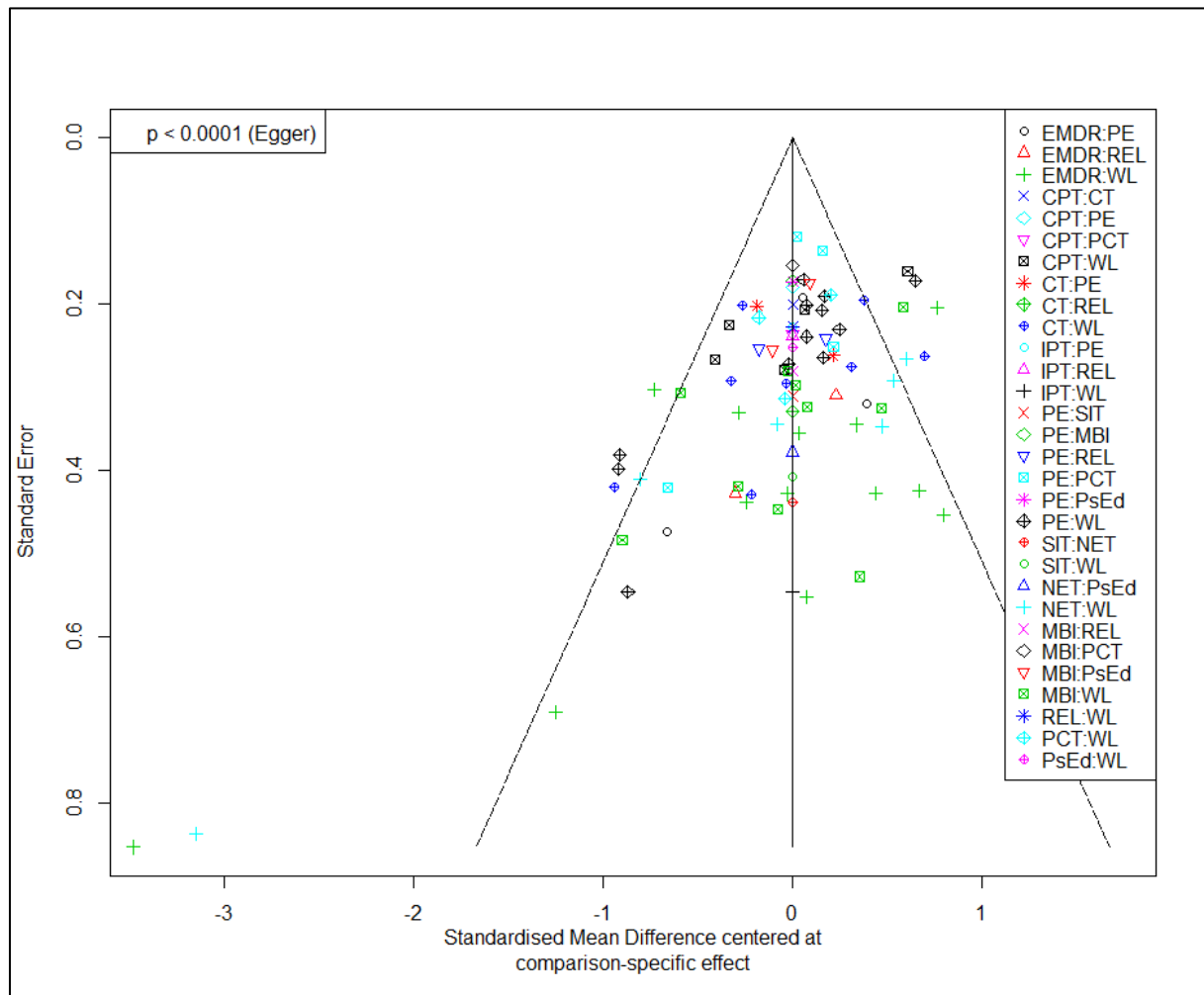
Appendix Q1

Results of the first run of the NMAoutlier algorithm for outlier detection: Plot displaying

Cook's Distance
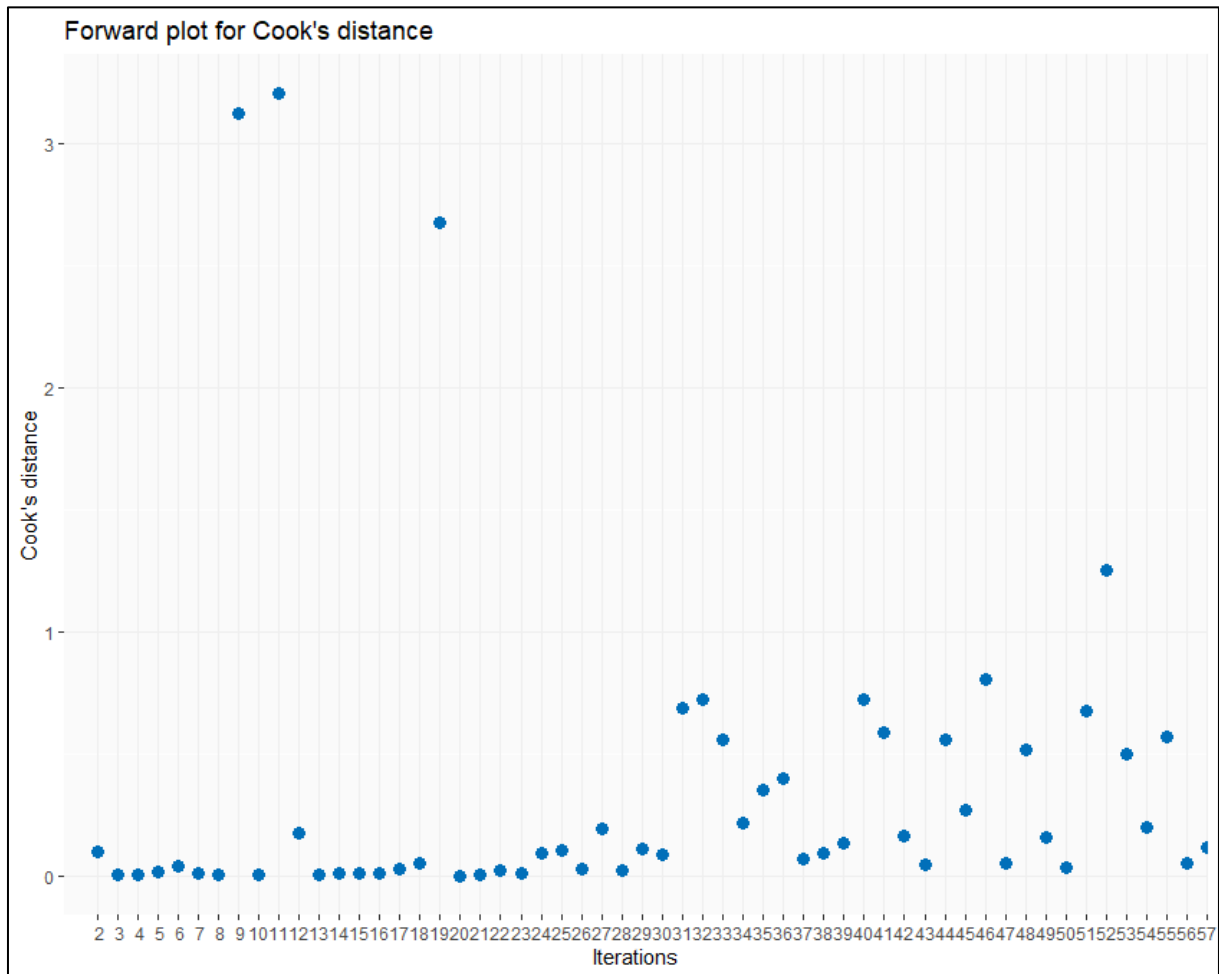
Note: Study entered in iteration 9 = Nidich et al., 2018; in iteration 11 = Markowitz et al.,

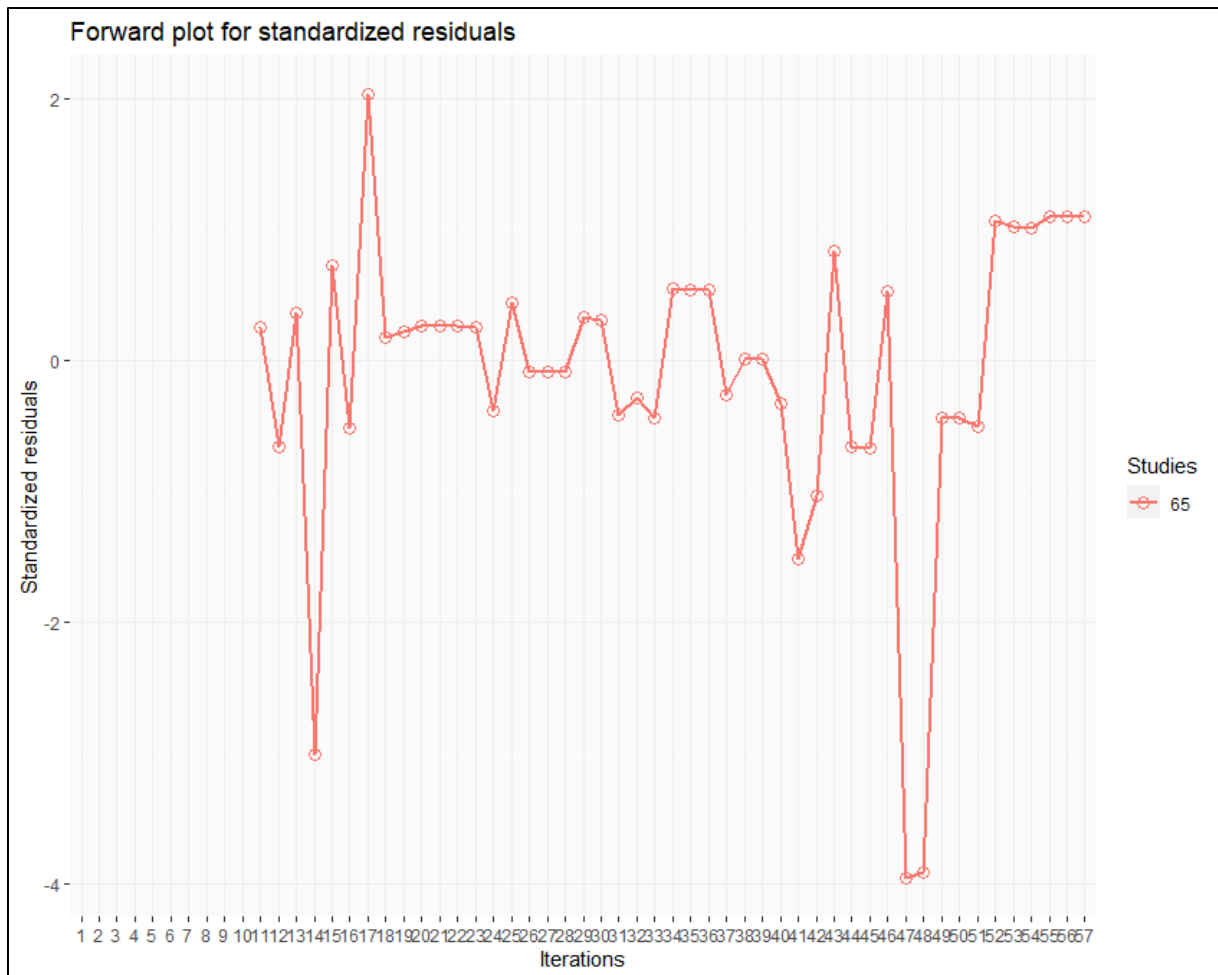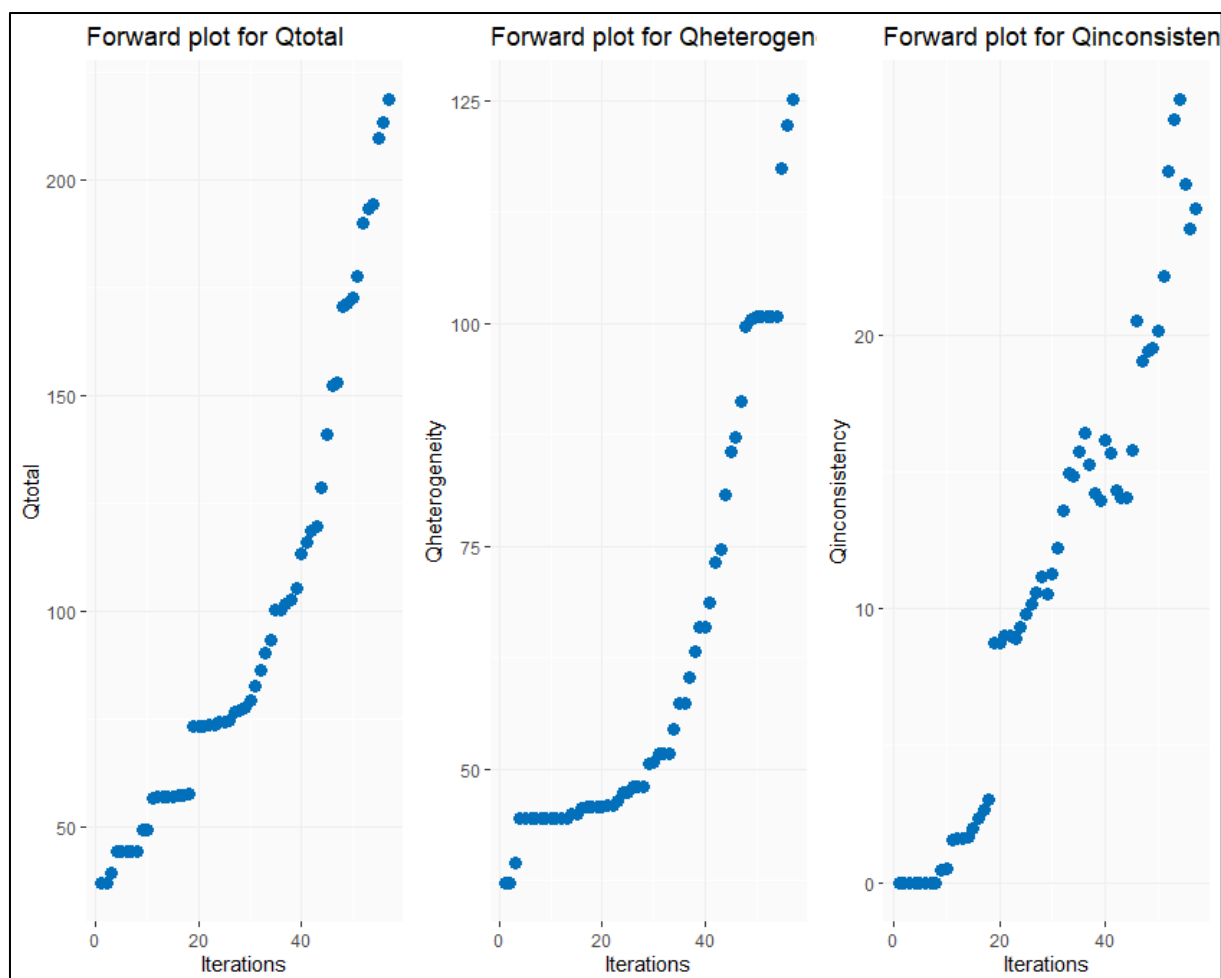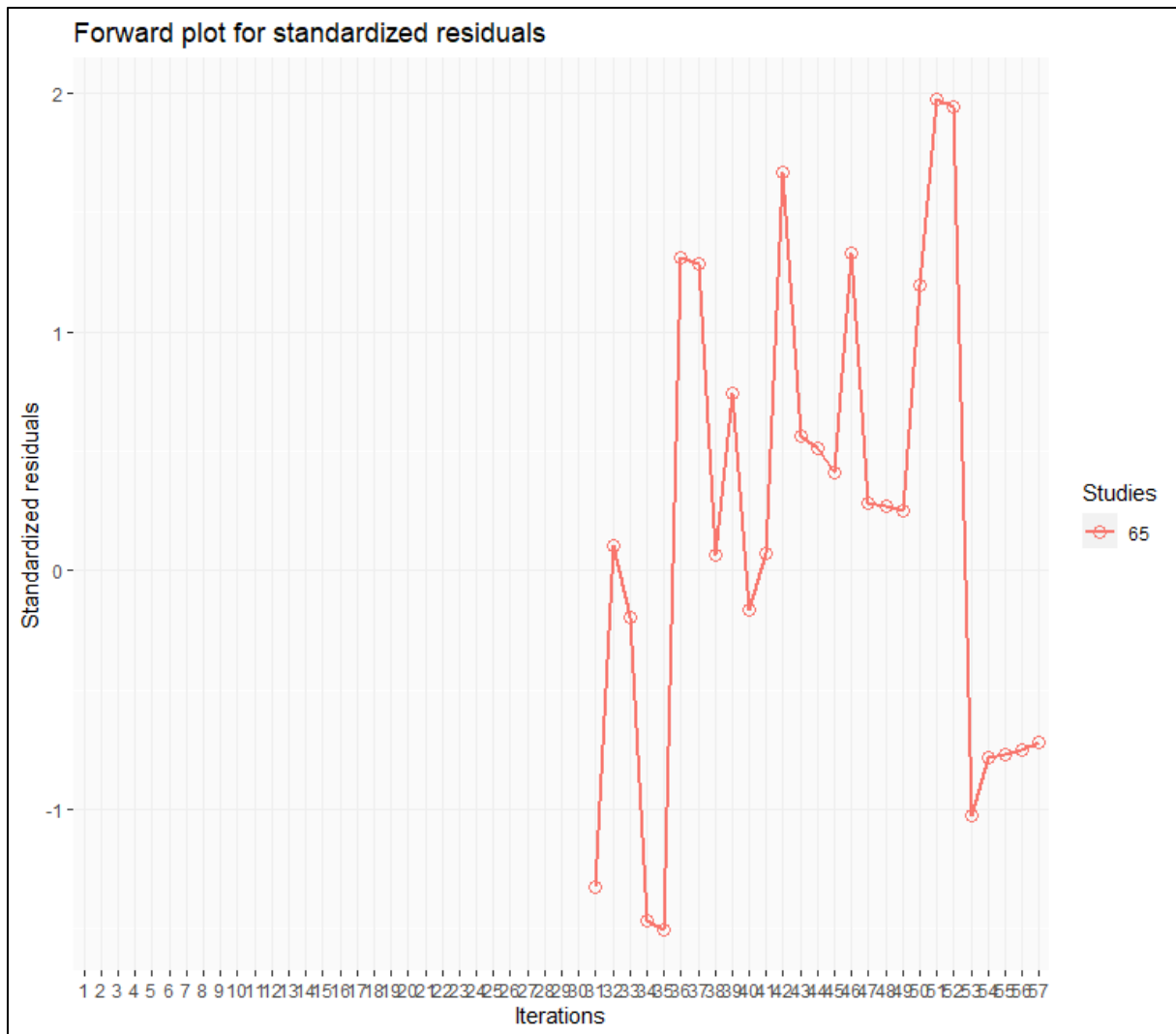2015; in iteration 19 = Rothbaum, Astin & Marsteller, 2005.

Appendix Q2

Results of the first run of the NMAoutlier algorithm for outlier detection: Plot displaying

standardized residuals for Nidich et al., 2018
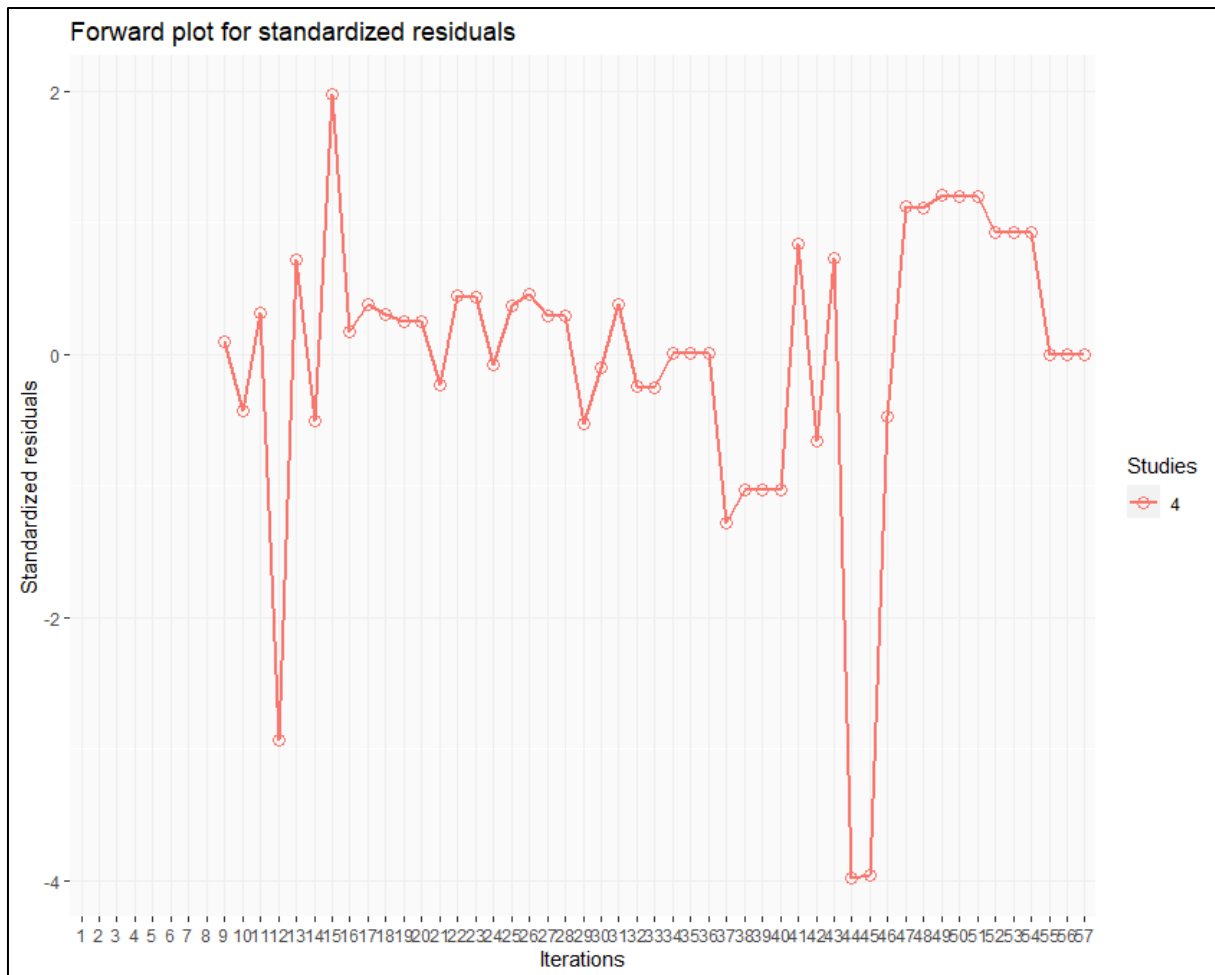
Appendix Q3

Results of the first run of the NMAoutlier algorithm for outlier detection: Plot displaying

standardized residuals for Markowitz et al., 2015

Appendix Q4

Results of the first run of the NMAoutlier algorithm for outlier detection: Plot displaying

standardized residuals for Rothbaum, Astin & Masteller, 2005
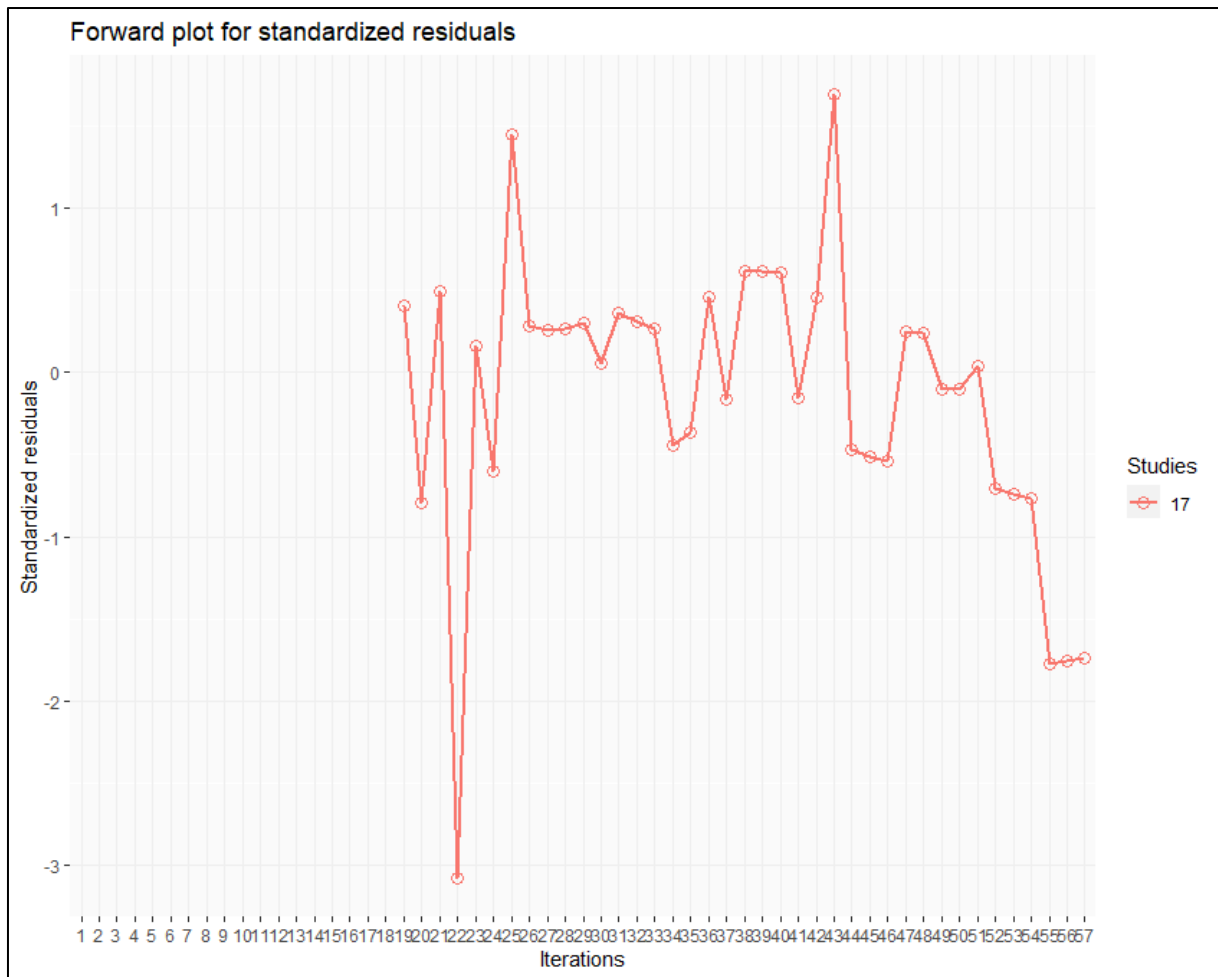
Appendix Q5

Results of the first run of the NMAoutlier algorithm for outlier detection: Plot displaying the

ratio of variances

Forward plot for ratio of variances
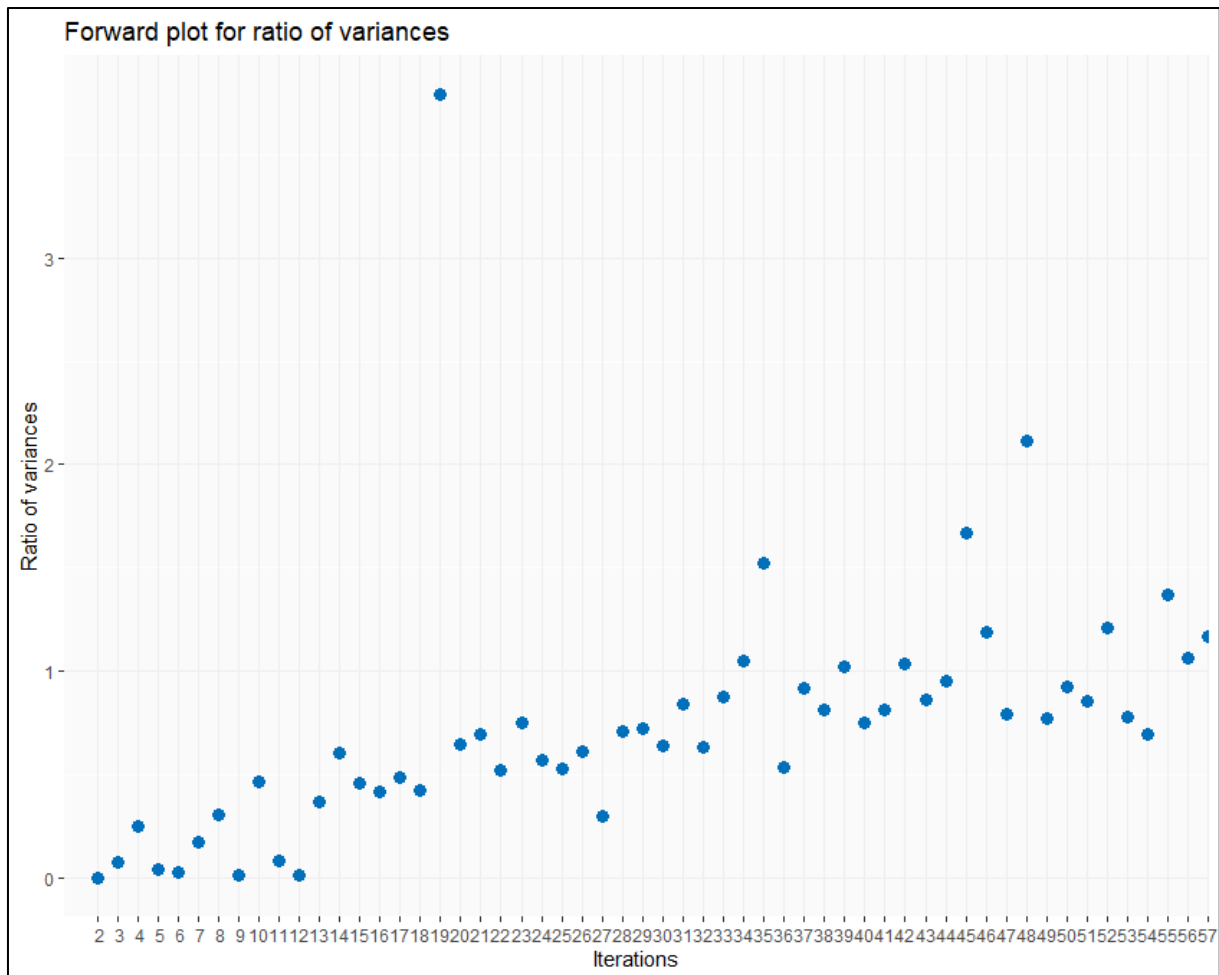
Note: Study entered in iteration 19 = Rothbaum, Astin & Marsteller, 2005.

Appendix Q6

Results of the first run of the NMAoutlier algorithm for outlier detection: Plot displaying

changes in heterogeneity and inconsistency (Q statistic)

Note: Study entered in iteration 19 = Rothbaum, Astin & Marsteller, 2005

Appendix Q7

Results of the second run of the NMAoutlier algorithm for outlier detection: Plot displaying

Cook's Distance
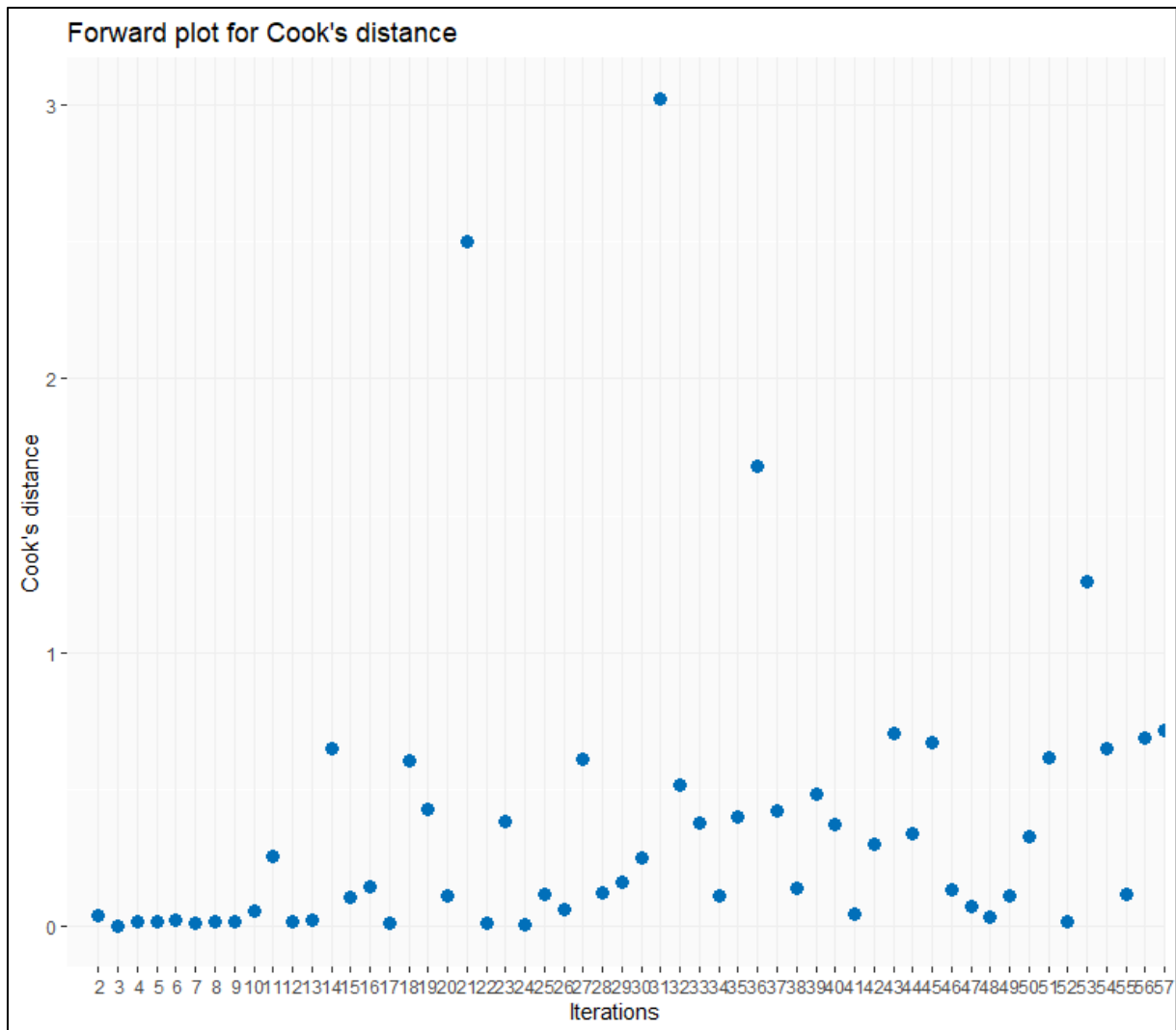
Note: Study entered in iteration 21 = Rothbaum, Astin & Marsteller, 2005; in iteration 31 =

Markowitz et al., 2015.

Appendix Q8

Results of the second run of the NMAoutlier algorithm for outlier detection: Plot displaying

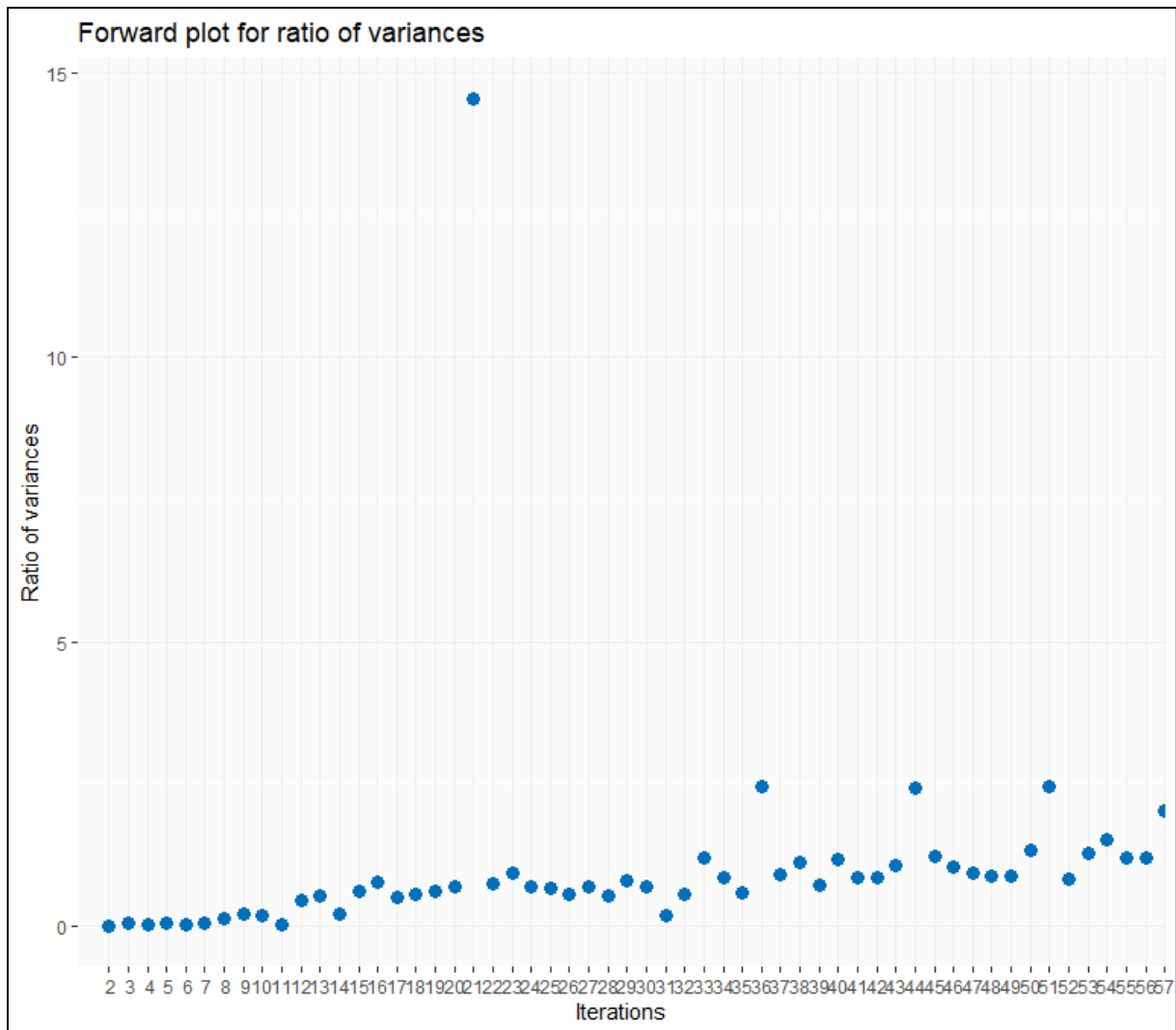standardized residuals for Markowitz et al., 2015

Appendix Q9

Results of the second run of the NMAoutlier algorithm for outlier detection: Plot displaying

standardized residuals for Rothbaum, Astin & Masteller, 2005
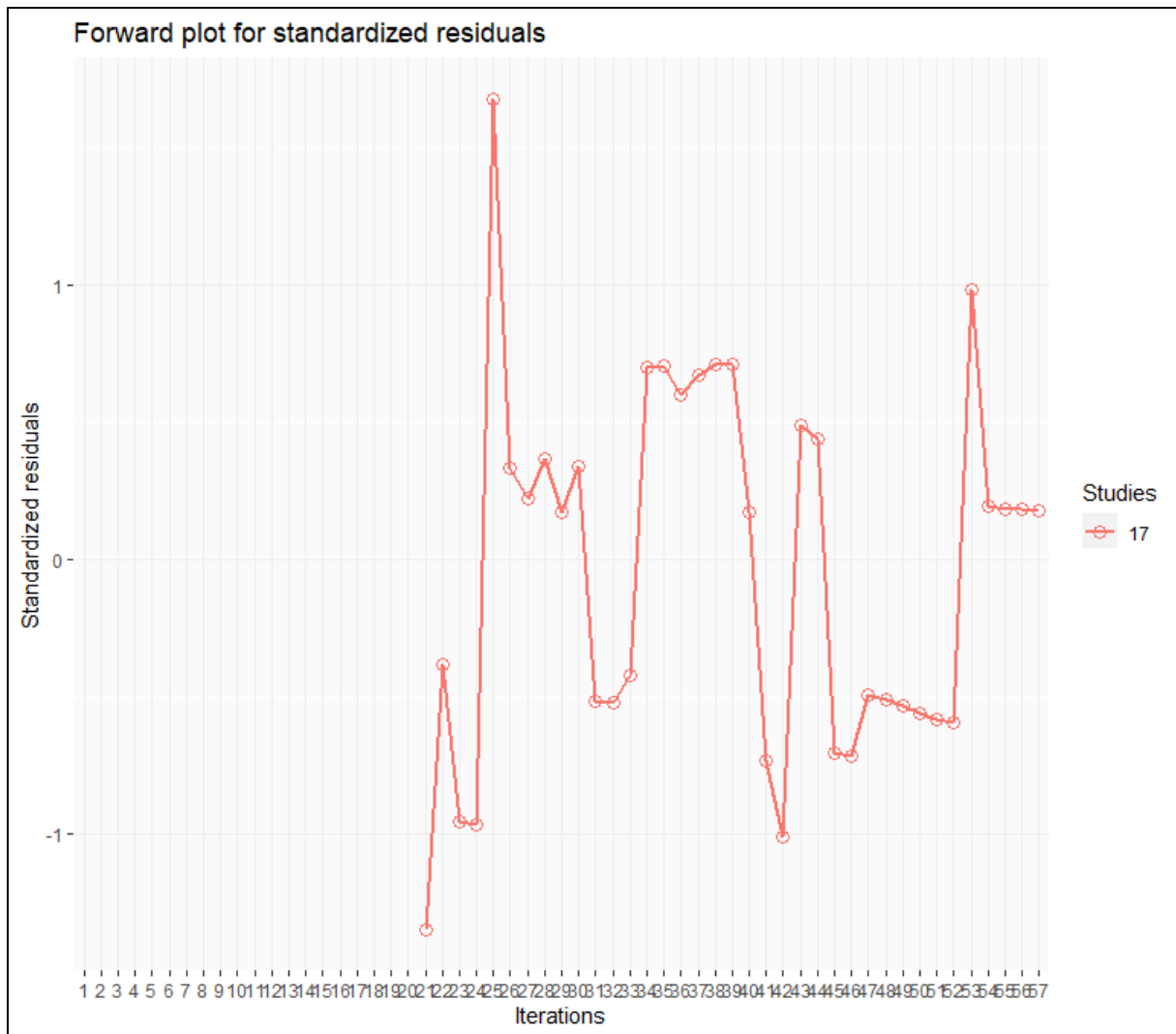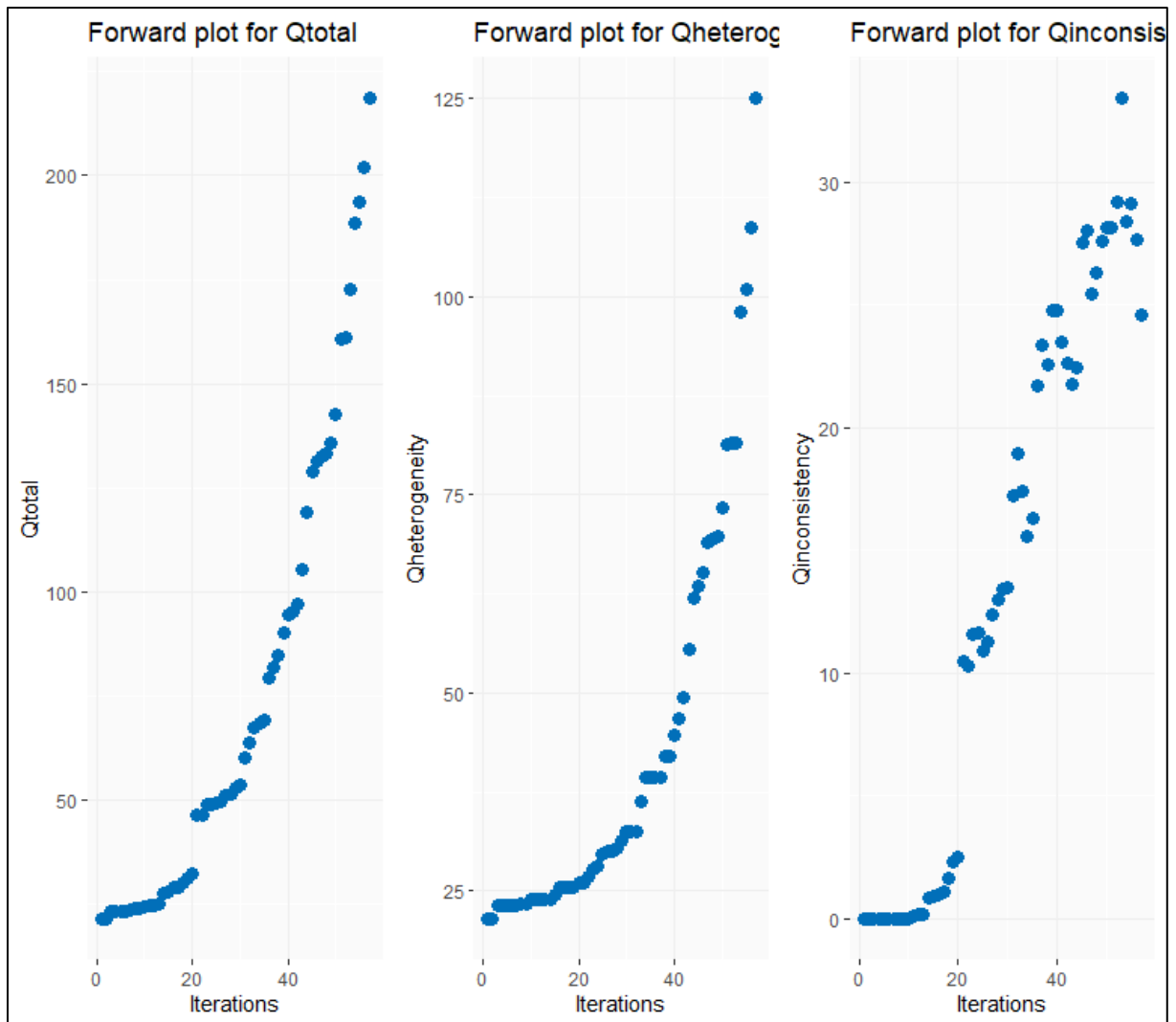
Appendix Q10

Results of the second run of the NMAoutlier algorithm for outlier detection: Plot displaying

the ratio of variances

Note: Study entered in iteration 21 = Rothbaum, Astin & Marsteller, 2005

Appendix Q11

Results of the second run of the NMAoutlier algorithm for outlier detection: Plot displaying

changes in heterogeneity and inconsistency (Q statistic)

Note: Study entered in iteration 21 = Rothbaum, Astin & Marsteller, 2005