

Accepted Manuscript

An empirical review of potential mediators and mechanisms of prolonged exposure therapy

Andrew A. Cooper, Erin G. Clifton, Norah C. Feeny

PII: S0272-7358(16)30288-4

DOI: doi: [10.1016/j.cpr.2017.07.003](https://doi.org/10.1016/j.cpr.2017.07.003)

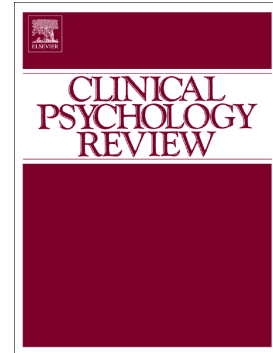
Reference: CPR 1617

To appear in: *Clinical Psychology Review*

Received date: 25 July 2016

Revised date: 27 April 2017

Accepted date: 9 July 2017



Please cite this article as: Andrew A. Cooper, Erin G. Clifton, Norah C. Feeny , An empirical review of potential mediators and mechanisms of prolonged exposure therapy, *Clinical Psychology Review* (2017), doi: [10.1016/j.cpr.2017.07.003](https://doi.org/10.1016/j.cpr.2017.07.003)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Running Head: MECHANISMS OF PROLONGED EXPOSURE

An Empirical Review of Potential Mediators and Mechanisms of Prolonged Exposure Therapy

Andrew A. Cooper^{1,2}

Erin G. Clifton²

Norah C. Feeny²

¹Department of Psychology, University of Toronto Scarborough

² Department of Psychological Sciences, Case Western Reserve University

Submitted: April 27, 2017

Preparation of this manuscript was supported by a grant to Dr. Feeny from the National Institute of Mental Health (R01 MH066348). The funding source had no involvement in any aspect of this manuscript.

Please address correspondence to Andrew A. Cooper, Department of Psychology, University of Toronto Scarborough. 1265 Military Trail, Toronto, Ontario Canada M1C 1A4. Email: aa.cooper@utoronto.ca Phone: 416-287-7092

Keywords: prolonged exposure; PTSD; mechanisms; inhibitory learning; belief change

Abstract

Prolonged exposure (PE) is an empirically-supported treatment for posttraumatic stress disorder (PTSD), but the precise mechanism(s) by which PE promotes symptom change are not well established. Understanding *how* PE works is critical to improving clinical outcomes, advancing dissemination efforts, and enhancing transdiagnostic models of psychopathology. However, mechanisms research conducted in clinical treatment settings is complex, and findings may be difficult to interpret without appropriate context. This is the first review of potential mechanisms of PE to provide such context, by rigorously evaluating empirical findings in line with essential criteria for effective research on mechanisms (or mediators). We begin by describing six putative mechanisms identified by emotional processing theory and contemporary models of fear extinction, before thoroughly reviewing empirical findings from clinical research on PE and similar PTSD treatments. We provide a detailed description of each study and mechanism test, as well as ratings of strength of evidence and quality of evaluation based on a novel rating scheme. We highlight variables with strong evidence (belief change and between-session habituation), intermediate evidence (inhibitory learning and emotional engagement), and minimal support (narrative organization and within-session habituation). After discussing limitations of the extant literature and this review, we summarize specific challenges for research on PE mechanisms and highlight directions for future study based on clinical and research implications.

An Empirical Review of Potential Mechanisms in Prolonged Exposure Therapy

Chronic posttraumatic stress disorder (PTSD) is a common, debilitating disorder associated with substantial symptom burden and impairment (see Cooper, Feeny & Rothbaum, 2015, for a recent review). There are several empirically-supported treatments for PTSD (Cusack et al., 2016), including prolonged exposure therapy (PE; e.g., Foa, Hembree, & Rothbaum, 2007), cognitive processing therapy (CPT; e.g., Resick & Schnicke, 1992), and cognitive therapy (CT; Ehlers et al., 2003). PE in particular has been designated as a first line treatment in many clinical guidelines (e.g. Institute of Medicine, 2008), achieving outcomes comparable to other trauma-focused treatments and superior to various control conditions across a variety of trauma types and populations (Powers, Halpern, Ferenschak, Gillihan & Foa, 2010). Yet despite a well-established theoretical basis and robust evidence of its efficacy, important questions remain with respect to PE's *mechanisms of change* – that is, the “active ingredients” of treatment that lead to and cause therapeutic improvement (Kazdin, 2007; Kindt, 2014). Research on mechanisms may involve different levels of measurement (e.g., behavioral, neurobiological) and is critical to the broader goal of identifying transdiagnostic processes and vulnerabilities linked to psychiatric impairments (e.g., RDoC; Cuthbert & Insel, 2013) and mechanisms shared across similar treatments. Identifying mechanisms of change may also help to optimize interventions by improving treatment response and lowering attrition (Kazdin, 2007) and may help advance dissemination efforts by addressing barriers to implementation and providers' concerns about adopting specific treatments.

Contemporary psychotherapy mechanism research typically focuses on the relationship between theoretically important change processes (e.g., acute changes in fear responding) and clinical outcomes (e.g., symptom improvement). While the term mechanism is ubiquitous in this area, most studies actually investigate *mediators*, which are interceding variables that statistically

account for the relationship between an intervention and outcome. Mediators can provide guidance about potential mechanisms but do not necessarily explain the cause of or reasons for change, and may in fact provide misleading or erroneous information. For this reason, Kazdin (2007, pp.5; see Table 1) proposed seven explicit criteria for evaluating mediators as part of a framework for investigating mechanisms. Unfortunately, there are a myriad of conceptual¹ and practical challenges posed by the study of mechanisms in clinical treatment samples, and few empirical studies meet the criteria proposed by Kazdin, an issue that often goes unmentioned in reviews of this type of research (for an exception, see Smits, Julian, Rosenfield & Powers, 2012).

The present paper offers an empirically-focused review of the literature, targeting processes that have received the greatest attention as potential mechanisms of PTSD symptom change for PE and similar exposure-based therapies. We focus on nominally psychological processes because of their dominant role in both theory and research on PE mechanisms, and the absence of an exhaustive and comprehensive review of this literature. To provide a more focused review of this vast and often complex literature, we do not extensively address theories based on other CBT variants for PTSD (e.g., Ehlers & Clark, 2000), other exposure-based therapies for anxiety (e.g., Mineka & Thomas, 2005), or neurobiological models of PTSD and its treatment with psychotherapy (e.g., Liberzon & Sripada, 2007; Kindt, 2014). We review empirical data relevant to six potential mechanisms identified by two dominant contemporary psychological theories relevant to PE: emotional processing theory (EPT; Foa & Kozak, 1986) and fear inhibition learning (e.g., Craske, Treanor, Conway, Zbozinek & Vervliet, 2014). For each putative mechanism, we evaluate the strength of evidence and methodological quality of their

¹ Conceptual questions include the appropriateness of linking change processes to specific treatment techniques (e.g., Doss, 2004), and the impact of patient and provider characteristics on the study of treatment processes (e.g., DeRubeis et al., 2014). These are important questions that largely fall outside the realm of this review.

constituent empirical studies. In the absence of a well-established metric for evaluating research on mechanisms (mediators), we developed an approach based on two related but separate concepts represented in Kazdin's seven key criteria (2007, Table 1). First, we characterize the *strength of evidence*; that is, the robustness and consistency of findings, both within and across studies. Second, we characterize the *quality of evaluation*, reflecting criteria that are intrinsically linked to aspects of study design, methods and analytic strategies. For example, one of Kazdin's key criteria is specificity, whereby a proposed mediator shows a single robust relationship between intervention and outcome, thus requiring a second candidate mediator for comparison purposes. Our ratings also incorporate other contemporary considerations related to overall study quality and risk of bias (e.g., Cusack et al., 2016), including representativeness, sample size, and handling of missing data. Readers are encouraged to review the Online Supplement to this article for further detail about the development of this approach and detailed ratings for each category. Summary scores for both subscales are listed alongside each empirical test are listed in Table 2. Finally, we summarize the current state of evidence for these mechanisms, concluding with a review of limitations and important directions for future research in this area.

Description of Prolonged Exposure Therapy

PE is a manualized cognitive-behavioral intervention for PTSD (e.g., Foa, Hembree, & Rothbaum, 2007). Treatment begins with collection of information about a patient's trauma history, including identification of a primary trauma which will be the focus of subsequent exposure activities. Early sessions involve psychoeducation about PTSD symptoms, common reactions to trauma, and the treatment rationale. Breathing retraining is taught as a form of relaxation. PE involves two exposure components: 1) confronting avoided trauma-related situations and reminders (i.e., *in vivo* exposure); and 2) repeatedly re-visiting the trauma memory

(i.e., imaginal exposure). *In vivo* exposures are based on a personalized hierarchy of trauma-related avoided, objectively safe, fear-provoking situations and scenarios (e.g., riding the bus, going to the grocery store, or crowded places). Patients repeat *in vivo* exercises multiple times as homework assignments between sessions, ideally remaining in previously avoided situations for sustained periods of time (i.e., 30 mins) or until their distress reduces. Imaginal exposure is the repeated recounting, or re-visiting, of the target trauma for a prolonged period of time in session. During imaginal exposure the therapist provides support and encouragement (e.g., “you’re doing a great job sticking with it”) and promotes emotional engagement with the trauma memory (e.g. encourages the inclusion of thoughts and feelings at the time). The therapist also facilitates emotional processing afterwards, which involves discussing the patient’s thoughts and feelings about the experience of recounting the trauma memory, providing support and normalizing reactions related to the trauma, and using open-ended questions to explore thoughts and feelings that may be contributing to the maintenance of PTSD. As between-session homework, patients are also instructed to listen daily to a recording of their imaginal exposure. PE is typically concluded after 8 to 15, 90-minute sessions when the patient’s PTSD symptoms have significantly reduced.

Contemporary Psychological Models of Prolonged Exposure

Emotional Processing Theory (EPT). EPT is a transdiagnostic theory that provides a general account of the causes and maintaining factors underlying PTSD and other fear-based anxiety disorders, as well as the core principles of successful exposure-based treatments (Foa & Kozak, 1986; Foa, Huppert & Cahill, 2006). Based on Lang’s bioinformational theory of fear (1977, 1979), EPT emphasizes the role of pathological fear structures within memory in which fear is represented as a cognitive structure that is a “blueprint” for escaping danger. This fear

structure includes representations of feared stimuli (e.g., a man with a gun), the fear responses (e.g., an increase in heart rate and sweating, running away), and the *meaning* associated with the stimuli (e.g., “guns are dangerous”; Foa & Kozak, 1986). A typical fear structure acts as a template for effective response to danger, whereas a pathological fear structure includes inaccurate associations and representations of the world, leading to links between objectively safe stimuli, fear, and escape/avoidance responses (Foa & Kozak, 1986). For example, a woman attacked while walking alone in a park at night might come to fear all parks and being alone at night, illustrating an erroneous part of the fear structure that may influence her behavior (e.g., avoiding parks). EPT also postulates that disrupted perceptual processing at the time of the trauma makes fear memories more disorganized and fragmented, enhancing the potential for erroneous associations and intrusive sensory experiences. The pathological fear structure of PTSD includes excessive stimulus and response elements, as well as pathological meaning elements (Foa et al., 2007) that influence emotional responding and promote avoidance of external and internal reminders of the trauma. Avoidance therefore serves as a critical maintaining factor. EPT also emphasizes cognitive factors related to development and maintenance of PTSD. Trauma-related beliefs are described as being related to inflexible, inaccurate perceptions of self, the world, others and one’s future; common examples include viewing the world as extremely dangerous, oneself as incapable of coping with stress or fear, and all others as being untrustworthy or intent on doing harm (Foa, Huppert, & Cahill, 2006). Such beliefs can develop anew in the wake of trauma, or exist prior to and be strengthened by the experience. EPT describes these cognitions as part of the pathological fear structures, and charges that they maintain PTSD by promoting avoidance (Foa & Kozak, 1986; Foa et al., 2006).

EPT emphasizes a number of potential mechanisms varying in the degree to which they

are linked to specific therapeutic activities or processes (e.g., Foa et al., 2007; Foa & McLean, 2015). Emotional processing and modification of the fear structure are described by Foa and Kozak (1986) as a multi-step process involving activation of the fear memory by elements that match part of the structure, also called *emotional engagement*, and provision of disconfirming evidence that is incompatible with the erroneous information. Repeated exposures to trauma reminders (i.e., *in vivo* exercises) and the trauma memory itself (i.e., imaginal exposure) promote *habituation* to fear. In learning terminology, this corresponds to the dissociation of stimulus from response elements or extinction of the fear response. The earliest iterations of EPT emphasized that this process could be measured by 1) reduced physiological activation, and 2) habituation both within and 3) between treatment sessions (Foa & Kozak, 1986). Repeated imaginal exposure practice also facilitates *re-organization of the trauma narrative*, which is theoretically linked in EPT to maintenance of intrusive symptoms and distress. EPT also describes *cognitive change* as a mediator of symptom improvements. Early iterations of EPT describe this process using learning theory terminology, focusing on both linguistic and non-linguistic “modification of meaning” and stimulus-stimulus associations (Foa & Kozak, 1986), such as the integration of a safe therapy context during imaginal exposure. A patient who has erroneous beliefs about danger inherent to a given situation may be able to state these beliefs (e.g., “Parks are never safe”) or may simply indicate them by their actions (e.g., avoiding all parks), with both processes reflecting cognitive change. Similarly, therapy tasks like *in vivo* or imaginal exposure can provide disconfirming evidence about probability and valence of beliefs via discussion or experiential processes. Contemporary descriptions of EPT (e.g., Foa et al., 2006; Foa & McLean, 2015) have particularly emphasized the mechanistic role of change in trauma-related beliefs, such as altering self-blaming or shameful views, with these processes less robustly linked to

specific exposure techniques. Other forms of cognitive change described in EPT include self-reflective processes such as a sense of mastery (e.g., Foa et al., 2007) or recognition of one's own distress tolerance abilities. Notably, EPT allows for overlap and interaction among change processes; for instance, habituation may facilitate cognitive change by disconfirming beliefs about anxiety spiraling out of control.

Fear Inhibition Learning Models. Fear inhibition models have received considerable research attention recently, perhaps because of their extensive translational and broad clinical literature base, or their suitability for integration with contemporary memory and neurobiological models (for reviews, see Kindt, 2014; Milad & Quirk, 2012). Based upon classic learning paradigms showing that conditioned fears can be easily extinguished but also easily recur (e.g., Bouton, 1993), these transdiagnostic models emphasize the mechanistic role of *fear inhibition learning* (Craske et al., 2014) underlying fear extinction in exposure therapies. PTSD results from the pairing of the traumatic experience (unconditioned stimulus) with sensory and contextual information (neutral stimulus), such that the presence of the previously-neutral stimuli indirectly activates memory of the trauma (conditioned stimulus; Bouton, 1993). This pairing is never fully erased. Instead, through repeated exposure to the conditioned stimulus in the absence of the trauma, a competing association is formed that signals the absence of fear or danger. Thus, this putative mechanism is analogous to the development of a separate competing memory, with the new association being strengthened by repeated exposures. Although they are distinct models, there is ample conceptual overlap and compatibility between this approach and EPT (e.g., Gillihan & Foa, 2011). For instance, emotional engagement with exposure exercises is important for inhibitory learning (e.g., Culver, Stoyanova & Craske, 2012), although this is emphasized in terms of preventing behaviors that impair learning processes, such as distraction

or reinforcement of avoidance (e.g., safety signals). However, with respect to some other mechanism variables highlighted by EPT, fear inhibitory models offer a divergent view. For example, although inhibitory learning may occur in parallel to habituation, reduction in distress or fear is not obligated to accompany strengthening of competing inhibitory information (Craske et al., 2014). Moreover, change in trauma-related beliefs is not central to these models (see Graham & Milad, 2011). Thus, while habituation and change in trauma-related beliefs may co-occur as a result of inhibitory learning or successful treatment, they are not strongly emphasized in these models.

Conceptual Clarifications. Our discussion of EPT and inhibitory learning models highlights two key areas in need of clarification for this review and the mechanism literature at large. First, the term “cognitive change” is widely used in theoretical and empirical papers to describe a variety of processes, resulting in ambiguity and definitional drift, both of which pose significant challenges to the study of key change processes. A myriad of cognitive variables are linked to PTSD, including negative trauma-related beliefs, rumination, threat-biased attention, general autobiographical memory, deficits in inhibitory learning and resistance to fear extinction (see Dalgleish, 2004; Ehling, Kleim & Ehlers, 2011; Jovanovic & Ressler, 2010; Milad & Quirk, 2012). These variables differ in terms of their temporal relationship to the occurrence of trauma and the onset of PTSD (i.e., in their designation as risk factors, concomitant processes, and/or maintaining factors), and their status as potential mechanisms in PTSD treatment models. They also vary in the degree to which they are “accessible” to individuals with PTSD, and by extension the ways in which they are capable of being assessed. In order to adequately evaluate cognitive change as a potential mechanism of PE, further delineation of these concepts is required. This issue is perhaps especially prominent in distinguishing between inhibitory learning

and changes in trauma-related beliefs described in EPT and other cognitively-oriented models (e.g., Ehlers & Clark, 2000). As such, we follow other authors (e.g., Hofmann, 2008; Mennin & Farach, 2007; Power & Dalgleish, 1997) in distinguishing between attitudinal and associative processes, describing these as *beliefs* and *inhibitory learning*, respectively. Beliefs are consciously-accessible appraisals and attitudes, such as negative views and judgments about oneself, abilities and future, expectations of harm and perceptions of control. These are higher-order, conceptual processes linked to rule-based learning (Power & Dalgleish, 1997). For example, changes in self-reported endorsement of the statement, “I am never safe in parks” would reflect belief change. In contrast, learning processes reflect aspects of classical conditioning, including changes in expectations and stimulus-response linkages. For the same example, inhibitory learning processes would underlie the development of a competing, alternative and neutral association between parks and personal safety. Per EPT, patients need not be consciously aware of propositional learning and changes to the fear structure (e.g., Foa & Kozak, 1986). These changes may be routinely described in language consistent with conditioning or learning paradigms (e.g., unconditioned stimulus expectancies), but are also patently cognitive processes, conscious or otherwise (Hofmann, 2008). EPT and other models (e.g., Ehlers & Clark, 2000) suggest that associative processes underlie or promote belief changes. Critically, the literature characterizes these not as strictly separate processes but potentially different levels of measurement, chiefly distinguished by their accessibility to patients’ awareness and thereafter, to self-report (see LeDoux, 2014).

A second, similar issue concerns use of the terms *habituation* and *extinction*. In learning theory, these concepts have distinct definitional boundaries: extinction refers to decreased conditioned responding (Bouton, 1993), whereas habituation is a simpler, non-associative

process by which repeated presentation of a stimulus leads to decreased responding. The role of habituation in experimental inhibitory conditioning paradigms is more complex and controversial (e.g., McSweeney & Swindell, 2002), but likely involves a hierarchical relationship in which non-associative processes occur in parallel to, and perhaps facilitate, extinction processes (Kamprath & Wotjak, 2004; Myers & Davis, 2007). Translational research paradigms (e.g., animal fear learning) allow for more precise investigation of the contributions of these processes, versus the more fluid context of clinical exposure (e.g., LeDoux, 2014). In practice, both terms are used extensively in contemporary exposure therapy theory and research (e.g., Gillihan & Foa, 2011) and both EPT and inhibitory learning are rooted in extinction research. However, the linkage between theoretical mechanisms and methods of measurement is somewhat more tenuous. As our review will show, this conceptual ambiguity occasionally results in the same measurement approach being described in contrasting terms. Unpacking potential mechanisms of change requires some definitional consistency, so for the purposes of this review, we primarily categorize mechanisms in line with authors' descriptions.

Psychological Mediators and Mechanisms of Exposure Therapy for PTSD

This review is focused on six processes: (1) *emotional engagement*, (2) *within-session habituation* and (3) *between-session habituation* to trauma-related fear, (4) *reorganization of the trauma narrative*, (5) *trauma-related belief change*, and (6) *inhibitory learning*. Notably, the relative importance of these variables has not been static over time. Revisions to EPT, particularly with regard to PTSD, have shifted the emphasis on some potential mechanisms, in line with updated clinical and empirical evidence. For instance, the role of trauma narrative change is given little attention in contemporary reviews (e.g., Foa & McLean, 2015) despite being the emphasis of considerable research a decade ago. We also recognize that controversies

may exist with respect to the role of these variables as true mechanisms of change or simply facilitative conditions. For instance, whereas habituation is described or classified as a mechanism by some sources (e.g., Asnaani, McLean & Foa, 2016; Gallagher & Resick, 2012; Rothbaum & Foa, 1997), others describe it simply as an indicator or evidence of emotional processing (e.g., McNally, 2007; Zalta, 2015). To be comprehensive, we review all variables that have consistently been implicated as potential mechanisms of change in PE despite shifts in emphasis over time (see Table 2 for a summary). Addressing conceptual questions about mechanism research is necessary for understanding the adequacy of our empirical evidence and ultimately uniting lines of research that may have been viewed as separate mechanisms up to this point. Furthermore, this approach may help clarify gaps in the literature and specific goals for future study on mechanisms.

Emotional Engagement. Emotional engagement is conceptualized as the activation of distress, fear, or anxiety during exposure techniques (e.g., Foa & Kozak, 1986; Foa et al., 2007). In PE, this process is most salient during in-session imaginal exposure to the trauma memory, and also relevant during *in vivo* exposure. Per EPT, emotional engagement is required to activate the fear structure, which in turn allows for modification and integration of new information. Under-engagement can occur under a number of conditions (Foa et al., 2007), including purposeful or unintentional avoidance efforts by the patient (e.g., distraction or speeding through upsetting content), and other emotional states such as dissociation or anger (Foa, Riggs, Massie, & Yarczower, 1995). Emotional engagement is often measured using the Subjective Units of Distress Scale (SUDS; Wolpe & Lazarus, 1966), a 0 to 100 scale with higher scores representing more severe distress. In typical PE protocols, patients provide SUDS ratings at the beginning, end, and at fixed intervals (e.g., every five or ten minutes) throughout in-session imaginal

exercises, helping to guide the exposure and monitor discomfort. Most clinical research in PE has used SUDS ratings (e.g., peak SUDS or change from baseline to peak) although engagement has also been measured behaviorally by rating facial fear (Foa, Riggs, et al., 1995), or using physiological proxies like heart rate (e.g., Pitman et al., 1996).

We identified five studies that evaluated the relationship between in-session emotional engagement and PTSD symptom improvement, with four studies finding evidence that emotional engagement was related to better treatment outcome in exposure-based treatments². However, findings were not robust or consistent across or even within studies, with considerable variability in how engagement was defined, when it was measured, and how outcomes were assessed, even among studies using SUDS ratings. An early, small open trial of PE demonstrated the numerically strongest link between engagement and outcomes, showing that greater reductions in post-treatment PTSD symptoms were significantly correlated with observer-coded facial fear ($r = .78$) and peak SUDS ($r = .71$) during imaginal exposure (Foa, Riggs, et al., 1995). By contrast, a recent study of women with PTSD and borderline personality disorder (BPD) who completed an open trial of PE (Harned, Ruork, Liu, & Tkachuk, 2015) reported no evidence linking engagement to loss of PTSD diagnosis at post treatment. Engagement was defined by pre-to-peak SUDS change, averaged across all imaginal exposures (i.e., in session and homework assignments), and also assessed for highest ratings of other emotions (i.e., fear, sadness, anger, guilt, shame & disgust) reported immediately before and after imaginal exposures. The other three studies of emotional engagement all reported mixed evidence linking engagement to outcomes. For instance, in Veterans receiving flooding therapy, change in resting-to-peak heart

² We omit Jaycox et al. (1998) from this section because their analyses of engagement do not reflect peak SUDS or change from baseline to peak SUDS, instead using average within-session habituation to derive cluster solutions in analyses we describe later.

rate during exposure was linked with experiencing fewer daily intrusive symptoms at post-treatment ($r = .70$; Pitman et al., 1996). However, several other psychophysiological measures of activation or arousal (e.g., heart rate, skin conductance) failed to predict symptom improvement; thus, the authors interpreted this as a negative finding with respect to the necessity of emotional engagement.³ In a study of women with interpersonal or sexual-violence related PTSD who received PE with or without cognitive restructuring, post-treatment PTSD symptoms were more robustly correlated with peak SUDS during the final imaginal exposure ($r = .48$) compared to the first exposure ($r = .09$; Rauch, Foa, Furr, & Filip, 2004). Yet another study found greater change from pre to peak SUDS during the first imaginal exposure among patients categorized as improved versus unimproved after PE (van Minnen & Hageraars, 2002). However, this study failed to find evidence that engagement during the second imaginal exposure, or during between-session homework assignments, differed between improved and unimproved patients. Moreover, across all PE completers, none of the engagement measures were significantly correlated with post-treatment PTSD symptoms after controlling for pre-treatment symptoms. Notably, this study found that patients who improved less robustly had higher anticipatory distress prior to initiating imaginal exposure for the first time, inconsistent with some previous findings suggesting anticipatory distress as a marker of better outcomes (Foa, Riggs, et al., 1995; Jaycox, Foa & Morral, 1998; Rauch et al., 2004).

In summary, these five studies highlight an inconsistent relationship between emotional engagement and symptom improvements in PTSD, with failures to replicate and method variance complicating straightforward interpretation of these results. None of these studies showed that emotional engagement during in-session imaginal exposure techniques was

³ Although flooding is a precursor to contemporary PE, the implementation of flooding and study design of Pitman et al (1996) is rather unlike other studies of imaginal exposure reported on in this review.

necessary to reduce symptoms, in contrast to how it is usually described in EPT (Gillihan & Foa, 2011) and implementation guides for PE (Foa et al., 2007). We suspect that establishing clearer guidelines for research on this variable might enhance our ability to elucidate its relationship to symptom change in PE. It may be prudent to (a) assess engagement at different intervals during treatment, accounting for potential differences between the initial and subsequent imaginal exposures, (b) control for baseline PTSD symptoms and anticipatory anxiety, as both of these have been found to correlate with emotional engagement (e.g., van Minnen & Hagedaars, 2006) and might serve as confounds; and (c) adopt observer ratings of emotions and behavior during exposure (e.g., Foa, Riggs, et al., 1995), which may improve precision in capturing behaviors and emotional states corresponding to engagement, as well as those which theoretically interfere, such as anger, distraction, or over-arousal (i.e., Foa et al., 2007).

Within-Session and Between-Session Habituation. In EPT, habituation is emphasized as a signal of modification of the fear structure, with diminished responding to trauma reminders or the trauma memory regarded as evidence of emotional processing (e.g., Foa & Kozak, 1986; Gillihan & Foa, 2011). In studies of PE, habituation has been measured on two timespans: *within-session*, typically assessed by the change from peak SUDS to final SUDS within a given imaginal exposure, and *between-session*, assessed by change in mean SUDS between sessions, with studies varying in terms of which sessions are compared. For reader clarity, we use the abbreviations WSH and BSH for this section.

Eight studies have evaluated WSH and BSH as predictors of PE outcome, with contrasting findings. There is very little evidence supporting WSH as a central mechanism of change in PE or among other exposure therapies for anxiety disorders (see Craske et al., 2008). Four studies assessed WSH during the first imaginal exposure, with three of these (Nacasch et

al., 2015; van Minnen & Foa, 2006; van Minnen & Hagedaars, 2002) also including a comparison to a later session. Most studies defined WSH as the difference between peak and final SUDS rating within an imaginal exposure, with one study using change in heart rate and other psychophysiological measures instead of SUDS (Pitman et al., 1996). None of these studies found a significant relationship between WSH at the first imaginal exposure and dependent variables reflecting superior treatment outcomes, with similar results for a study that used averaged WSH scores across imaginal exposures (Harned et al., 2015). Two additional studies using more complex analytic strategies, including cluster analysis and HLM also failed to find evidence linking WSH to treatment outcomes (Jaycox et al., 1998; Sripada & Rauch, 2015). In contrast, a recent study using mixed models found that greater WSH was associated with more robust symptom improvement at the next session of PE, with greater WSH across treatment also linked to more rapid improvement and better overall response (de Kleine, Smits, Hendriks, Becker & van Minnen, 2015). Notably, two studies in the aforementioned group compared conventional 90-minute PE to a shorter 60-minute protocol, with a resulting decrease in time spent conducting imaginal exposure (Nacasch et al., 2015; van Minnen & Foa, 2006). Both studies observed greater WSH in the longer duration treatment protocol, and provided figures depicting a drop-off in SUDS levels in the final stages of each imaginal exposure. This suggests a critical need to consider duration of imaginal exposure in studies evaluating mean and peak-post change in SUDS as potential mechanisms.

In contrast to WSH, several studies have found evidence that BSH is associated with superior outcomes in PE or similar treatments for PTSD. Twelve studies have tested BSH as a predictor of superior outcomes across samples varied in trauma and demographic characteristics (e.g., Veterans, sexual assault victims, women with comorbid borderline personality disorder).

Five studies measured BSH by comparing peak or mean SUDS from the first exposure to the final exposure (Gallagher & Resick, 2012; Harned et al., 2015; Nacasch et al., 2015; Rauch et al., 2015; van Minnen & Foa, 2006). Each of these studies reported modest but significant associations between BSH and outcome of PE, including change in PTSD symptoms across treatment ($r = .24$; Gallagher & Resick, 2012; $r = .40$, Nacasch et al., 2015; $r = .36$, Rauch et al., 2004), post treatment PTSD symptoms ($r = .30$; van Minnen & Foa, 2006), and loss of PTSD diagnosis ($\eta^2 = .39$; Harned et al., 2015). Two studies evaluated mean between-session change in peak SUDS (deKleine et al., 2015; Rothbaum et al., 2014). One study linked this variable to greater change over time and lower post-treatment PTSD-symptoms in PE-treated patients (deKleine et al., 2015), whereas the other found a link between this variable and improved outcome in just one of three exposure therapy conditions (Rothbaum et al., 2014). A psychophysiological measure of BSH (i.e., change in resting-to-peak heart rate from first to final session) was moderately correlated with post-treatment PTSD symptoms in Veterans who received flooding therapy ($r = .51$; Pitman et al., 1996). Another study focused on BSH between the first and second imaginal exposure session, comparing outcomes between improved and unimproved patients (van Minnen & Hageraars, 2002). Improved patients showed greater BSH, and after partialling out baseline PTSD symptoms, between-session habituation was moderately correlated with post-treatment symptoms ($r = .36$). Finally, a recent study assessed reliable change criteria in both mean and peak SUDS from the first and final PE session, noting that relatively few patients experienced this degree of improvement (28% and 35% for mean and peak SUDS, respectively; Bluett, Feeny & Zoellner, 2014). Patients with reliable change in SUDS had lower PTSD and depression symptoms and generally superior functioning at post-treatment. However, there were no differences in post-treatment PTSD diagnostic status on the

basis of reliable change. The authors interpreted this finding as indicating BSH is not strictly necessary for improvement, and suggested that some patients might instead be learning distress tolerance skills (e.g., Craske et al., 2008). Finally, two studies examined patterns of change in SUDS ratings across time and clinical outcomes in PE. A cluster analysis of mean SUDS ratings across six sessions of PE identified three response patterns: (1) high initial SUDS with gradual decline over sessions; (2) high SUDS with minimal decline over sessions, and; (3) moderate initial SUDS with minimal change over sessions (Jaycox et al., 1998). All three groups evidenced symptom improvements, but patients in the first group had superior outcomes, with all analyses also controlling for duration of exposure. In an open PE trial for Veterans with PTSD, hierarchical linear modeling (HLM) was used to evaluate patterns of SUDS change across treatment, using multiple SUDS ratings within each imaginal exposure, with robust estimation of missing data points (Sripada & Rauch, 2015). Post-treatment PTSD symptom change and treatment responder status were both predictive of between-session SUDS change, with comparable results found in analyses of treatment completers. Both of these studies represent substantial advances in methodological sophistication, particularly with respect to making efficient use of all available data, but are also limited somewhat by small sample sizes.

In summary, the strength of evidence for BSH as a mechanism of PE differs substantially from that of WSH. Our brief review of the latter variable reflects the limited empirical data supporting its relation to outcome in PE as well as other exposure treatments. As noted, WSH appears to be more robustly associated with the duration of imaginal exposures (Nacasch et al., 2015; van Minnen & Foa, 2002), suggesting the importance of accounting for duration of imaginal exposure procedures in studies of this variable. By contrast, multiple research groups have presented evidence for BSH, across a range of populations and analytic strategies including

some that account for temporal precedence and missing data, with some concordance in measurement strategies. All twelve studies we reviewed found at least partial evidence that this variable was related to superior outcomes, with some studies using multiple assessments across treatment (e.g., DeKleine et al., 2015; Jaycox et al., 1998; Sripada & Rauch, 2015). The three studies failing to find consistent relations were unlike other studies in terms of exposure implementation (i.e., repeated flooding procedures; Pitman et al., 1996; comparison of extinction facilitators; Rothbaum et al., 2014) and method of measuring habituation (e.g., reliable change; Bluett et al., 2014). Yet most studies have reported modest relationships between BSH and clinical outcomes, and it is therefore possible that this process may be sufficient but not necessary for symptom improvement (e.g., Bluett et al., 2014), or co-occurring with other change processes. Establishing consensus guidelines for how best to operationalize BSH, including *when* it is assessed, may prove critical in future studies, particularly for efforts to link to neurobiological processes or transdiagnostic research on exposure therapy. A multi-modal assessment may provide critical in comparing patients' subjective ratings (i.e., SUDS) to psychophysiological metrics (i.e., HR) as predictors of change.

Organization of the Trauma Narrative. Many PTSD theories suggest that traumatic memories are disorganized and more fragmented than memories of non-traumatic events, including problems with the sequence of events or missing “chunks” of time with no organic cause (e.g., Foa et al., 2006; Halligan, Michael, Clark, & Ehlers, 2003). In PE, memory organization is thought to be facilitated by repeated imaginal exposures, in-session processing of the trauma memory with the therapist, and between-session homework assignments that involve listening to audiotapes of the memory (Foa, et al., 2007). The majority of research in this area has focused on changes to organization of oral or written narratives of trauma memories, under

the premise that these correspond to changes in the trauma memory itself. Two major reviews of the PTSD trauma narrative literature have concluded that inconsistent measurement and terminology impairs the ability to draw strong conclusions about ties between trauma, PTSD and indices of memory disruption (O’Kearney & Perrott, 2006; Bedard-Gilligan & Zoellner, 2012). This limitation also affects the ability to integrate findings with contemporary cognitive and neurobiological models of memory (e.g., Tronson & Taylor, 2007), perhaps explaining its omission from recent reviews of potential mechanisms (e.g., Foa & McLean, 2015; Zalta, 2015).

Three studies have tested the relationship between narrative change and PTSD symptoms after treatment with PE (Bedard-Gilligan, Zoellner & Feeny, in press; Foa, Molnar & Cashman, 1995; van Minnen, Wessel, Dijkstra & Roelofs, 2002). The earliest of these studies used a novel coding scheme to compare *fragmentation* (repetitions, unfinished thoughts, and speech fillers) and *organization* (utterances reflecting realization, decision-making or planning) between the first and final imaginal exposure (Foa, Molnar et al., 1995). In this sample, final narratives were longer and more organized, with change in fragmentation correlated with change in PTSD ($r = .73$). A subsequent study using the same narrative coding system did not strongly support these findings, although direct comparison was complicated by their focus on improved versus non-improved patients (van Minnen et al., 2002). The authors of this follow-up speculated that improved organization and decreased fragmentation might be general side effects of PE as opposed to essential mechanisms. A recent study used a comprehensive coding method to rate narratives representing a positive memory, a negative memory, and patients’ trauma memory, comparing pre- and post-treatment versions in patients treated with PE or sertraline (Bedard-Gilligan et al., in press). Pre-treatment trauma narratives were significantly more fragmented than control narratives, but fragmentation indices were highly associated across narrative types,

suggesting the influence of patients' general recounting style. Moreover, narrative change was not robustly linked to outcome, and was comparable across PE and sertraline despite the relative lack of focus on retelling and processing the trauma memory in the latter treatment.

In summary, there is little evidence for trauma narrative reorganization as a central mechanism of PE. In addition to the underwhelming empirical data, several reviews have questioned if narrative changes reflect underlying enduring organizational modifications to the memory, or are more closely linked to factors related to repeated retelling, such as anxiety associated with the initial disclosure (Bedard-Gilligan & Zoellner, 2012). Post-imaginal exposure processing in PE may lead to therapists' feedback or interpretations being incorporated into the narrative structure by the client. Further, narrative organization may simply be a by-product of certain treatment procedures, or recovery from PTSD in general. Findings reported by Bedard-Gilligan and colleagues (in press) further suggest that patient narrative "style" may play the largest role in the structure of the post-treatment trauma narrative. In light of this poor evidence base and relative decline in, the three studies described in this section (i.e., Bedard-Gilligan et al., in press; Foa, Molnar et al., 1995; van Minnen et al., 2002) were not evaluated using the rating scheme developed for this review.

Belief Change. Belief variables include self-perceptions of enhanced mastery and distress tolerance ability, with negative trauma-related beliefs having received far and away the most empirical attention and emphasis in theoretical models (Ehlers & Clark, 2000; Foa et al., 2006). The most commonly used measure of these negative beliefs is the Post-Traumatic Cognitions Inventory (PTCI; Foa, Ehlers, Clark, Tolin, & Orsillo, 1999), a 36-item self-report of negative cognitions associated with PTSD including self-blame, negative beliefs about oneself, and negative beliefs about the world. These beliefs are empirically and theoretically linked to

PTSD, a fact that has been supported by comparison to individuals with other disorders (e.g., Kleim, Ehlers & Glucksman, 2012). Other authors have classified trauma-related beliefs as members of a broader class of cognitive variables dubbed *threat reappraisals* which have been found to mediate symptom improvements across CBT treatments of a variety of anxiety disorders (Smits et al., 2012). EPT emphasizes that these negative beliefs maintain PTSD, as well as undermining efforts at recovery and maintaining avoidance behaviors (Foa et al., 2006; Foa & McLean, 2015). Furthermore, EPT postulates that reductions in these beliefs underlie successful recovery from PTSD, regardless of the method by which this is achieved (i.e., PE, another treatment or natural recovery).

To date, eight studies have examined the relationship between negative trauma-related beliefs and symptom change in PE. Three studies found that PTCI change from pre- to post-treatment was significantly associated with concurrent PTSD symptom change ($r = .60$, Hagens, van Minnen & de Rooij, 2010; $r = .58$, Foa & Rauch, 2004; $r = .41$, Nacasch et al., 2015). One of these studies also showed that change in PTCI across treatment predicted post treatment PTSD symptoms, but this effect was not maintained after controlling for change in PTSD symptoms during treatment. Another study also showed that PTCI change was uncorrelated with between-session habituation ($r = .01$) and both were trend level predictors of PTSD symptom improvement, which the authors interpreted as evidence that cognitive change and between-session habituation might be separate processes or indicators of improvement (Nacasch et al., 2015). Five recent studies have examined the temporal relation between belief change and PTSD symptom change in exposure therapies for adults with PTSD (Cooper, Zoellner, Roy-Byrne, Mavissakalian & Feeny, in press; Kumpula et al., 2016; McLean, Su & Foa, 2015; Økstedalen, Hoffart & Formo-Langkaas, 2015; Zalta et al., 2014). These studies used

various robust statistical methods (e.g., lagged mixed effects regression) to estimate the magnitude of association between belief change and PTSD improvement while accounting for missing data, controlling for the influence of prior symptom levels, and testing temporal sequence (i.e., is belief change a predictor or product of symptom change?). All five studies found evidence that belief change significantly predicted subsequent PTSD symptom improvement in PE treatments. Three studies also convincingly demonstrated a unidirectional relationship, with belief change preceding symptom change more robustly than the reverse relationship (Cooper et al., in press; Øktedalen et al., 2015; Zalta et al., 2014), with the remaining studies suggesting a more equivocal relationship (Kumpula et al., 2016; McLean et al., 2015). Several of these studies also included important comparisons between PE and other treatments in terms of this relationship. One study showed a significant moderating effect of treatment on the association between belief and symptom change, with a much stronger relationship in patients treated with PE ($d = 0.93$) versus those treated with sertraline ($d = 0.35$; Cooper et al., in press). Another study of Norwegian inpatients compared PE with and without imagery rescripting (IR), focusing on change in guilty and shameful thoughts. The authors found that within-patient changes in guilt and shame predicted subsequent PTSD improvements in both protocols. Finally, a recent study of patients treated with one of four treatments for PTSD and comorbid alcohol use disorder (i.e., permutations of PE or supportive counseling, with naltrexone or placebo) found no difference in the strength of the relationship between negative belief and PTSD change between three of four treatment groups, with weaker effects found in patients who received supportive counseling and placebo (McLean et al., 2015)⁴. Thus, even

⁴ Belief change has been shown to mediate subsequent PTSD improvement in adolescent girls with sexual assault-related PTSD randomized to 8-14 weekly sessions of PE for adolescents or client centered therapy (McLean, Yeh, Rosenfeld, & Foa, 2015).

those with PTSD and co-morbid alcohol use disorder showed a relationship between negative beliefs and PTSD symptom change.

In summary, negative trauma-related belief change processes have strong evidence as mechanisms of PE. All eight studies we reviewed found evidence of a predictive relationship to PTSD symptom change, with several studies (e.g., Zalta et al., 2014) clearly establishing a temporal pattern consistent with the theoretical mechanism effect, and use of a common measure (i.e., PTCI) facilitating comparisons between studies. Evidence for this mechanism would be strengthened by including other ways of measuring negative beliefs (e.g., via behavioral tasks or independent assessors) as well as more studies comparing the strength of these relationships across modalities. Testing other theoretically important appraisal-type variables, such as subjective sense of mastery or ability to tolerate distress, would also provide important information about specific kinds of belief change that drive symptom improvement, and perhaps shed light on the relationship between belief change and cognitive learning processes. This is perhaps especially important in light of changes to PTSD in DSM-5 (APA, 2013), whereby cognitive alterations are now a core diagnostic symptom. Researchers will have to confront the issue of whether their investigations of trauma-related belief change must now be framed as sequential symptom change analyses, and how to account for overlap between typical measures of trauma related beliefs (i.e., PTCI) and content capture by DSM5-adapted measures of PTSD symptoms.

Fear Inhibitory Learning. There is strong evidence that extinction learning underlies successful exposure-like paradigms that reduce conditioned fear in animals and analogue samples (Bowers & Ressler, 2015; Jovanovic et al., 2010; Norrholm et al., 2011). However, there are numerous challenges to evaluating change in associative learning processes in PE,

especially in the context of clinical treatment research (i.e., Kindt, 2014; LeDoux, 2014). A particular difficulty is how best to establish that improvements are mediated by extinction versus other co-occurring processes, such as between-session habituation or trauma-related belief change. Several chemical compounds have been found to enhance inhibitory learning processes across a variety of animal and translational populations, as well as in a growing number of clinical samples (for more detailed reviews, see Burton et al., 2015). In general, these *cognitive enhancers* are administered just before the beginning of exposure procedures, in order to facilitate extinction. In exposure therapies for PTSD, the clearest empirical evidence of fear inhibition learning stems from studies that evaluate the effects of cognitive enhancers on response to exposure. When augmented conditions outperform control conditions, this is viewed as evidence in support of fear inhibition (or extinction) learning as a primary driver of symptom change. Critically, this comparison hinges on the augmentation effect being sufficiently strong to detect an advantage for cognitive enhancer conditions. Although a variety of compounds have been tested, D-cycloserine (DCS) and hydrocortisone have received the greatest attention across a variety of disorders, including PTSD (Burton et al., 2015; de Quervain, Aerni, Schelling & Roozendaal, 2009). DCS is a partial *N*-methyl-D-aspartate (NMDA) agonist that is thought to enhance fear extinction through its effects on NMDA glutamatergic receptors in the basolateral amygdala. Hydrocortisone, a glucocorticoid, is thought to facilitate extinction learning processes via similar glutamatergic pathways, and also suggested to improve tolerability of emotional distress during exposure techniques (deQuervain et al., 2009).

We identified four published and two unpublished studies⁵ involving randomized comparisons of DCS to placebo in PE or similar exposure therapies for PTSD. Results were

⁵ We include these two studies in our review, but omit them from Table 1, due to lack of detail on specific outcomes and sample characteristics.

mixed, with just one published study (Difede et al., 2013) clearly finding a superior outcome on primary PTSD measures for patients randomized to DCS augmentation versus placebo. Two as yet unpublished DCS augmentation trials were reported by Burton et al. (2015). In one small study of patients with full or sub-threshold PTSD randomized to DCS or placebo, DCS patients showed faster response and greater improvements in PTSD, depression, anxiety, and SUDS but experienced a return of symptoms at 3-month follow up (Henn-Haase et al., 2010). Another study tested DCS augmentation versus placebo prior to imaginal or *in vivo* exposure in an exposure-heavy CBT protocol for PTSD (Guay, March, & Landry, 2010). No differences were found in PTSD symptoms at post-treatment or 6 month follow-up between patients who received DCS or placebo. Two published studies compared DCS to placebo augmentation of treatments involving imaginal exposure. A study of Veterans who received a six-session treatment (four imaginal exposures) unexpectedly found worse outcomes for the DCS-augmented conditions (Litz et al., 2012). Average change in terms of clinician-assessed and self-reported PTSD symptoms actually constituted a worsening of symptoms, with fewer responders to DCS versus placebo at post-treatment (30% versus 70%, respectively). Post-hoc comparison of SUDS ratings suggested that DCS may have inadvertently strengthened the fear memory following the first imaginal exposure. Subsequent research in exposure paradigms for other disorders has supported this contention (e.g., Smits et al., 2013), and also shown that giving DCS post-exposure is not a viable remedy for this issue (Tart et al., 2013). Another study found no evidence of greater symptom response in patients who received DCS versus placebo (deKleine, Hendriks, Kusters, Broekman, & van Minnen, 2012). DCS-treated patients had higher rates of response in both intent-to-treat and completer analyses, but neither model controlled for significantly higher baseline symptoms in the placebo group. A recent secondary analysis of this data (de Kleine et

al., 2015) failed to find evidence consistent with greater extinction in DCS versus placebo treatment using a range of different possible extinction metrics. For instance, after controlling for initial SUDS and baseline PTSD severity, final SUDS ratings predicted next-session PTSD symptoms, but this relationship did not differ between DCS and placebo (i.e., no moderation). Similar patterns were observed for change in peak SUDS between sessions (e.g., between-session habituation) and change from peak to post SUDS within session (e.g., within-session habituation). Notably, the use of these metrics as a proxy for inhibitory learning is somewhat controversial, given that change in subjective ratings of distress is not required for fear learning to occur (Craske et al., 2014).

Finally, two studies of DCS augmentation used virtual reality-augmented exposure treatments with mixed to positive results (Difede et al., 2013; Rothbaum et al., 2014). Although their protocols differed slightly, in both cases patients completed in-session imaginal exposures while also viewing simulated images related to their trauma (i.e., 9/11 World Trade Center attacks or a virtual combat environment for Veterans of conflicts in Iraq and Afghanistan). One study found that patients who received DCS showed greater symptom improvements versus those in the placebo condition at post-treatment and 6-month follow-up and had higher rates of PTSD remission (Difede et al., 2013). The second study found no overall differences in PTSD reduction between patients who received DCS, placebo, or alprazolam augmentation, across treatment or the 12-month follow-up period (Rothbaum et al., 2014). However, average between-session change in peak SUDS was significantly predictive of post-treatment PTSD scores only for the DCS treatment condition, which the authors interpreted as evidence supporting its effect in augmenting inhibitory learning. Similarly, among patients who completed a standardized startle assessment, only patients in the DCS condition showed significantly attenuated response

after treatment; however, this analysis was only conducted in a relatively small subset of the sample (Rothbaum et al., 2014). In one small study of augmentation with hydrocortisone, patients who received augmentation were more likely to complete treatment and to attend more sessions, accounting for superior outcomes relative to placebo (Yehuda et al., 2015). Consistent with hypotheses, changes in glucocorticoid receptor sensitivity were highly correlated with symptom improvements ($r = -.86$) in the augmentation condition. Furthermore, lifetime PTSD severity and higher baseline glucocorticoid sensitivity were markers of superior response to augmentation, suggesting potential markers of superior response to treatment enhancement.

In summary, there is no doubt that the comprehensive theoretical and translational research base for inhibitory learning makes a strong case for its important role in PE (e.g., Bowers & Ressler, 2015; Jovanovic et al., 2010; Norrholm et al., 2011). However, evidence from augmentation studies involving PE has been underwhelming in terms of establishing a clear relationship between cognitive enhancers (e.g., DCS), greater inhibitory learning, and superior outcomes.⁶ The most rigorous and thorough tests of a mechanistic relationship (e.g., DeKleine et al., 2015; Rothbaum et al., 2014) have yielded mixed results. At this time, there is no well-established way of assessing inhibitory learning directly in clinical treatment contexts, nor is it clear that such variables would be representative of patients' subjective experience of improvements, a fact highlighted as a potential shortcoming of this model by some of its proponents (Kindt, 2014; LeDoux, 2014). However, we view as promising the fact that some contemporary studies (Rothbaum et al., 2014) have sought to incorporate secondary measures of inhibitory learning (e.g., startle paradigms). Nevertheless, while we remain optimistic about the

⁶ Tests of between-treatment effects (e.g., RCT designs) were substantially more common for inhibitory learning than any other mechanism. While this format is methodologically rigorous, it may be somewhat more difficult to demonstrate a strong effect in line with the rating scheme used in this review.

future of fear inhibitory models of PE and other exposure treatments, we must acknowledge the present limitations of this account.

Limitations of the Current Literature & this Review

Our review provides the most extensive contemporary summary of empirical evidence for psychological mechanisms of PE, which we hope will facilitate future research in this area and transdiagnostic comparison to other treatments and conditions. Spanning more than 25 years, our review includes many important and influential studies that shaped clinical and theoretical work on PE. These studies have paved the way for increasingly nuanced examinations of how exposure-based treatments work. However, it must also be noted that most fail to meet criteria for contemporary mechanism research described by Kazdin (2007). Indeed, our concern about how methodological quality might affect interpretation of results was what led us to develop a rating scheme for the empirical tests included in this review. On average, quality of evidence ratings were quite low, with a median score of just 4 ($M = 4.7$, $SD = 2.9$) of a possible 14 points. As detailed in the Online Supplement, few studies involved a test of specificity, gradient, temporal precedence, or even a formal test of mediation. Similarly, studies varied substantially in the management of missing data and attrition, albeit often in line with changes in dominant statistical approaches across research epochs. However, this is particularly relevant in terms of the adequacy of “dose” of PE, raising questions such as whether patients who complete a single exposure should be compared to those who complete treatment (e.g., Harned et al., 2015). Notably, the mean effective sample size for primary mechanism tests was 61.0 ($SD = 43.3$) with roughly a third of studies having fewer than 30 PE-treated patients. On the plus side, most studies involved heterogeneous trauma types and mixed genders ($k = 9$), and ratings of representativeness suggest findings may be broadly applicability. The relatively poor showing

with respect to quality of evidence does not seem to be limited to research on PE and PTSD in particular. For instance, in a targeted review of threat appraisal as a transdiagnostic mediator of change in cognitive behavioral treatments for anxiety disorders, Smits and colleagues (2012) noted that few studies involved designs that could conclusively demonstrate a causal relationship, with just over half incorporating a formal test of mediation. Indeed, a similar picture has emerged in mechanism-oriented reviews for a number of different treatments and diagnoses (e.g., Romano & Peters, 2015; van der Velden et al., 2015).

One challenge presented to readers of this and similar reviews stems from method variance in assessment, design, and analyses across the empirical studies highlighted in this collection. These studies involve a range of different approaches that are not always intuitively comparable to one another, nor inherently hierarchical. For example, studies ostensibly evaluating the same construct (between-session habituation) sometimes involve a single time point (e.g., during the first imaginal exposure; van Minnen & Hagedaars, 2002) versus averaging across all similar ratings (e.g., Jaycox et al., 1998) or constructing a multi-level model to compare effects across time (e.g., Sripada & Rauch, 2016). Yet even upon considering our quality of evaluation ratings and related literature (i.e., Kazdin, 2007), the best methods and approaches are not always apparent and are instead empirical questions in their own right. While acknowledging this as a limitation of the current literature, this issue also represents an opportunity for growth and development of this research domain, as highlighted in the final section of this manuscript.

This review is limited in a number of ways. First, because of the relatively small literature base, we did not focus exclusively on randomized trials. Similarly, we chose to represent statistical results as the authors did, eschewing detailed discussion of potentially

spurious findings driven by multiple tests or smaller samples. Second, in the interest of providing a focused review, we did not extensively discuss relevant findings from other variants of exposure therapy, or similar anxiety and fear based disorders. There are excellent studies (e.g., Wisco, Sloan, & Marx, 2016) and reviews (e.g., Smits et al., 2012) in other areas of research that may provide useful insights into the key change processes that underlie similar treatments or techniques. We focused on mechanisms identified as central to EPT and extinction models, omitting other variables that have been suggested as potentially important change processes in exposure based treatments for PTSD. Promising targets include development of skills such as distress tolerance (e.g., Bluett et al., 2014) or improved emotion regulation abilities (e.g., Cloitre, Stovall-McClough, Miranda, & Chemtob, 2004), as well as specific therapeutic activities that may proxy other change processes (e.g., homework compliance; Cooper, Kline, Graham, Bedard-Gilligan, Mello, Feeny & Zoellner, 2016). Transdiagnostic change processes may also prove important, such as the role of sleep quality in relation to treatment outcome (e.g., Resit, Gory & Hollifield, 2017). Finally, the utility of our rating scheme is as yet untested. This metric was designed as a straightforward and interpretable approach to provide quantifiable information about the quality of information and robustness of findings in an empirical mechanism test. We welcome feedback from readers on alternative methods for assessing these questions, and call upon mechanism researchers to develop consensus guidelines for such an approach.

Conceptual Challenges to Studying PE Mechanisms in Clinical Contexts

We now turn to broader conceptual considerations relevant to the question of how PE works, many of which are critical challenges for mechanism research in general (e.g., DeRubeis, Gelfand, German, Fournier & Forand, 2014; Kazdin, 2007; Smits et al., 2012). Most of the studies in our review explore mechanism-outcome relations with the implicit assumption that

these are stable across patients, providers, and time. Yet there are theoretical and empirical reasons to be dubious of this assumption. For illustrative purposes, consider two hypothetical clients: an employed, 50-year old man with no psychiatric history reporting primarily avoidance symptoms after a car accident, and a 21-year old woman with severe PTSD and comorbid depression tied to recurrent childhood sexual abuse. Clinicians may find themselves instinctively envisioning the unique challenges presented by each of these cases, and forming hypotheses about how best to implement PE to address such issues. Even if both patients receive PE that is implemented using a structured protocol (e.g., in a clinical trial), the content of their individual treatments is likely to vary, perhaps substantially, based on the characteristics and behaviors of both the patient and provider. For instance, clinicians may intuit that *in vivo* exposure is relatively more important for the first client, or may be more cautious about initiating imaginal exposure for the severely depressed client, as suggested by some research on therapist implementation of PE (van Minnen, Hendriks & Olf, 2010). Tailoring treatment to meet the needs of the client is a key element of PE (e.g., Foa et al. 2007), which is efficacious for a variety of trauma types, in diverse populations, with complex clinical comorbidities (e.g., Bedard-Gilligan et al., 2015). Nevertheless, the fact that PE *works* across these contexts does not necessarily mean that it *works the same way* for every patient. Critically, the active, flexible and transactional nature of therapy, which may account for PE's robust efficacy, also presents several significant challenges to mechanism research that must be addressed in future studies.

First, there is a need for greater consideration of on how patients' pre-treatment characteristics (e.g., PTSD symptoms, trauma type, demographics, comorbidities and psychological characteristics) may relate to outcomes in general, and also interact with potential therapeutic mechanisms. Individual differences can predict or predispose patients to achieve

better or worse clinical outcomes, although research on broad demographic and diagnostic variables has not yielded strong candidates in PE (e.g., van Minnen, Arntz & Keijsers, 2002). In the extremes, pre-treatment traits may predict clinical outcomes so robustly as to effectively negate the variance accounted for by kind or quality of treatment received for some patients, thereby reducing the magnitude of any true mechanism-outcome relationship that might be in effect (DeRubeis et al., 2014). Likewise, pre-treatment characteristics may also predict differential response to one treatment over another, suggesting a “prescriptive” relationship to achieve better outcomes (Murphy, Cooper, Hollon & Fairburn, 2009). Insofar as PE may achieve effects via multiple different mechanisms, particular patient traits may suggest a better or worse “match” for one mechanism (e.g., highly negative beliefs and attitudinal change) or technique (e.g., avoidance and *in vivo* homework). The incidence of these traits could either enhance or diminish the observed relationship between a given mechanism and outcomes, potentially leading to failure to replicate effects across studies. Testing patient characteristics as predictors of outcome, and controlling for such differences (e.g., propensity scoring) may help improve precision of mechanism analyses. Similarly, theory-driven relationships could be evaluated first in analyses testing moderation (e.g., Moser, Cahill, & Foa, 2010) or mediation (e.g., Clifton, Feeny, & Zoellner, 2017), and later via experimental research designs. As with patient characteristics, therapist-related variables may be related to treatment outcomes. Adherence to treatment guidelines and competence in treatment delivery have been linked to clinical outcomes in CBTs (e.g., Branson, Shafran & Myles, 2015) and may also impact the variance attributable to specific treatment mechanisms. For instance, it is possible that skillful PE therapists would be more capable of responding to challenging clients or tailoring treatment exercises to meet their needs. Likewise, a less experienced therapist might more timidly

implement imaginal exposures, resulting in insufficient patient engagement or processing (Foa et al., 2007). This adds a further layer of complication when considering the temporal relationships implied by a therapist making a “course correction” in how PE is implemented. In such a case, failing to conduct multiple assessments or consider non-linear relationships might lead researchers to inadvertently miss a true process-outcome relationship.

Although some mechanisms are linked to specific treatment activities in EPT (e.g., engagement and imaginal exposure), the ways in which these processes may overlap and interact are not clearly delineated. Change processes may occur simultaneously or sequentially, or may simply constitute different levels of measurement of the same underlying processes. Just as the impact of a given mechanism may vary *between* patients, it is plausible that mechanisms may be more or less important at different times *within* a given patient’s treatment (i.e., early versus late sessions). This concept is exemplified by research on sudden gains in PE, with several studies showing that patients who experience such rapid between-session symptom improvements ultimately achieve superior clinical outcomes (e.g., Jun, Zoellner & Feeny, 2013). While it is possible to simultaneously evaluate between- and within-patient variance in outcomes associated with candidate mechanism variables (e.g., Curran & Bauer, 2007), analyses of this sort are largely missing from our review (for an exception see Øktedalen et al., 2015).

Finally, the studies we review almost exclusively focus on predictors of PTSD symptom improvements across acute treatment, in line with the dominant model of assessing the therapeutic efficacy. However, this approach ignores other important variables affected by PTSD, such as social and occupational functioning, quality-of-life, and improvements in health and other psychological symptoms (see Goldfried, 2015), which also improve after PE treatment (Cusack et al., 2016). While it is possible that these sorts of improvements may be indirect

products of symptom abatement, this is an empirical question as yet untested in the literature. In a similar vein, it is unclear whether the mechanisms that drive symptom improvement are the same ones responsible for sustained remission or relapse prevention (e.g., Holtzheimer & Mayberg, 2010). For instance, between-session habituation might drive PTSD symptom reduction but be unrelated to ways of dealing with re-emergence of fear and other symptoms.

Future Directions and Recommendations

Mechanism research may receive an important boost from the RDoC research paradigm (e.g., Cuthbert & Insel, 2013) and accompanying shifts in federal funding priorities, a sentiment echoed in a recent Institute of Medicine opinion on improving clinical outcomes (Pincus & England, 2015). Reduced emphasis on randomized clinical trials may facilitate a renewed emphasis on fundamental processes underlying recovery, whether these occur naturalistically or are active ingredients of treatment (Goldfried, 2015). However, there are also challenges in the implementation of this framework within the field, including what to do with existing diagnostic labels and the role of replication and clinical utility (Zoellner & Foa, 2016). Perhaps a corresponding increase in access to archival datasets will facilitate secondary analyses of mechanism questions. Based on our review of this literature, we highlight three likely frontiers of advancement for the study of PE mechanisms in the coming years, and offer related recommendations for research design in Table 3.

First, enhanced measures of mechanism and outcome variables will be especially critical to the effort to identify processes of change in PE. Researchers should strive to incorporate multimodal assessment techniques, including observer ratings, psychophysiological metrics and neurobiological assays (i.e., biomarkers). Increasing access to and familiarity with technology in the form of smartphones and other personal tech devices should expand the cost-efficient use of

experience sampling techniques and physiological measurements (e.g., Walz, Nauta & Rot, 2014). These technologies also facilitate greater ability to assess clinical outcomes beyond symptom improvement, including factors linked to relapse prevention. Similarly, increasingly powerful statistical techniques (e.g., latent growth mixture modeling) offer the potential to conduct more complex analyses, making full use of longitudinal data even if it is incomplete (Tasca & Gallop, 2009). These approaches facilitate examination of key conceptual issues, such as the relative importance of between- and within-patient variance, and are available in most common statistical packages (e.g., SAS, SPSS, Mplus, R). Critically, when these sorts of approaches are not possible, researchers can enhance conventional analyses (e.g., regression) by attending to fundamental principles of mechanism research (Kazdin, 2007), such as establishing temporal precedence of predictors.

Second, the next wave of PE mechanism research may increase focus on specific therapeutic techniques. Future gains in this area will come from bottom-up translational approaches, informed by both clinical research and comprehensive psychobiological models of PTSD (e.g., Bowers & Ressler, 2015). These frameworks can provide insights into change processes that are difficult to study in clinical treatment samples and inform smaller-scale experimental modifications of how PE is implemented. For instance, animal research on fear inhibition can yield concrete, testable strategies to enhance impact of individual exposures (Craske et al., 2014). Critically, basic science and analogue studies are only part of a comprehensive translational approach (e.g., Ehring et al., 2011). The most scientifically rigorous designs should test alternative hypotheses (e.g., other candidate mechanism variables) and also evaluate if findings are similar in treatments other than PE. It may be especially important to account for treatment implementation variables, such as duration of imaginal exposure. As recent

studies demonstrate, these can shift from nuisance variables to the focus of clinical interest (e.g., Nacasch et al., 2015).

Finally, we expect that there will be increased interest in intra-individual differences, in terms of both pathways and predictors of treatment response. Heeding calls for a more personalized approach to medicine (e.g., Pincus & England, 2015), this research may primarily focus on identification of prescriptive variables that “match” patients to specific interventions. There are a host of theoretically relevant variables for study, with symptom clusters being obvious primary candidates given the diagnostic heterogeneity of the diagnosis (e.g., Galatzer-Levy & Bryant, 2013). Bolstered by its emphasis in the RDoC framework, neurobiological research may also prove especially important in this domain. Biomarkers have been found to predict development of PTSD (Michopolous, Norrholm & Jovanovic, 2015) and response to PE (Yehuda et al., 2013). By extension, this line of research may one day help to identify those for whom exposure therapy is likely to be most effective. Indeed, two studies have already reported patient characteristics associated with superior response to exposure augmentation with DCS (de Kleine, Hendricks, Smits, Broekman & van Minnen, 2014) and hydrocortisone (Yehuda et al., 2013). Smaller scale methodological changes may prove especially important for the future of PE mechanism research, including controlling for structural components of treatment (e.g., duration of imaginal exposure) as noted above. However, in some cases, the distinction between control variable and potential moderator is subtle, and requires careful attention to both theory and empirical data. For example, while controlling for anticipatory anxiety might improve precision in estimates of engagement during imaginal exposure, this variable might also reflect an intrinsic patient trait worthy of investigation as a moderator (e.g., distress tolerance).

Future studies could also attend to the ways in which patient and provider traits may

attenuate process-outcome relationships (e.g., DeRubeis et al., 2014), perhaps by use of propensity scoring of prognostic predictors of outcome. An ancillary benefit of this research frontier is that secondary analyses using cutting-edge analytic techniques may extend the impact of previously published RCT data. Data aggregation methods, including meta-analyses and so-called “mega analyses” (e.g., Jayawickreme et al., 2014), also hold the potential to identify predictors of response across groups of smaller samples, improving power to detect moderate effects. We hope that researchers might agree to uniform measures and increased data sharing practices to achieve these goals, and offer some suggestions in Table 3. Ultimately, advances in the understanding of *how* PE works should enhance clinical outcomes as well, by streamlining treatment, reducing attrition, and enhancing its appeal to patients and providers alike.

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Asnaani, A., McLean, C.P., & Foa, E.B. Updating Watson & Marks (1971): How has our understanding of the mechanisms of extinction learning evolved and where is our field going next? *Behavior Therapy*, 47, 654-668. doi: 10.1016/j.beth.2016.02.003
- Bedard-Gilligan, M., Jakob, D., Jeanne, M., Doane, L. S., Jaeger, J., Eftekhari, A., ... & Zoellner, L. A. (2015). An investigation of depression, trauma history, and symptom severity in individuals enrolled in a treatment trial for chronic PTSD. *Journal of Clinical Psychology*, 71, 725-740. doi: 10.1002/jclp.22163
- Bedard-Gilligan, M., & Zoellner, L. A. (2012). Dissociation and memory fragmentation in post-traumatic stress disorder: An evaluation of the dissociative encoding hypothesis. *Memory*, 20(3), 277-299. doi: 10.1080/09658211.2012.655747
- Bedard-Gilligan, M., Zoellner, L.A., & Feeny, N.C. (in press). Is trauma memory special? Trauma narrative fragmentation in PTSD: Effects of treatment and response. *Clinical Psychological Science*.
- Bluett, E. J., Zoellner, L. A., & Feeny, N. C. (2014). Does change in distress matter? Mechanisms of change in prolonged exposure for PTSD. *Journal of Behavior Therapy and Experimental Psychiatry*, 45(1), 97-104. doi: 10.1016/j.jbtep.2013.09.003
- Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin*, 114, 80-99. doi: 10.1037/0033-2909.114.1.80
- Bowers, M. E., & Ressler, K. J. (2015). An overview of translationally-informed treatments for posttraumatic stress disorder: Animal models of Pavlovian fear conditioning to human

- clinical trials. *Biological Psychiatry*, 78, E15-E27. doi:10.1016/j.biopsych.2015.06.008
- Branson, A., Shafran, R., & Myles, P. (2015). Investigating the relationship between competence and patient outcome with CBT. *Behaviour Research and Therapy*, 68, 19-26. doi:10.1016/j.brat.2015.03.002
- Burton, M.S., Youngner, C.G., McCarthy, A.J., Rothbaum, A.O. & Rothbaum, B.O. (2014). Enhancing exposure therapy for PTSD using d-cycloserine. In: Safir, M., Wallach, H. & Rizzo, A.A. (Eds.). *Future Directions in Posttraumatic Stress Disorder: Prevention, Diagnosis and Treatment*. (pp.347-365). New York, Springer.
- Clifton, E.G., Feeny, N.C., & Zoellner, L.A. (2017). Anger and guilt in treatment for chronic posttraumatic stress disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, 54, 9-16.
- Cloitre, M., Stovall-McClough, K. C., Miranda, R., & Chemtob, C. M. (2004). Therapeutic alliance, negative mood regulation, and treatment outcome in child abuse-related posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology*, 72(3), 411-416. doi:10.1037/0022-006X.72.3.411
- Cooper, A.A., Feeny, N.C., & Rothbaum, B.O. (2015). Clinical aspects of trauma-related anxiety and PTSD. In Pine, D., Ressler, K., & Rothbaum, B.O. (Eds.) *Primer on anxiety disorders: Translational perspectives on diagnosis and treatment* (pp. 135- 152). New York: Oxford University Press.
- Cooper, A.A., Kline, A.C., Graham, B., Bedard-Gilligan, M., Mello, P.G., Feeny, N.C. & Zoellner, L.A. (2017). Homework “dose,” type, and helpfulness as predictors of clinical outcomes in prolonged exposure for PTSD. *Behavior Therapy*, 48, 182-194. doi:10.1016/j.beth.2016.02.013

Cooper, A.A., Zoellner, L.A., Roy-Byrne, P., Mavissakalian, M.R., & Feeny, N.C. (in press).

Comparing the temporal relation of belief and symptom change in Prolonged Exposure and sertraline treatment of PTSD. *Journal of Clinical and Consulting Psychology*.

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy*, 58, 10-23. doi:10.1016/j.brat.2014.04.006

Craske, M. G., Kircanski, K., Zelikowsky, M., Mystkowski, J., Chowdhury, N., & Baker, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour Research and Therapy*, 46, 5-27. doi:10.1016/j.brat.2007.10.003.

Culver, N. C., Stoyanova, M., & Craske, M. G. (2012). Emotional variability and sustained arousal during exposure. *Journal of Behavior Therapy and Experimental Psychiatry*, 43, 787-793. doi:10.1016/j.jbtep.2011.10.009

Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, 62, 583-619. doi:10.1146/annurev.psych.093008.100356

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. doi:10.1016/j.cpr.2015.10.003

Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Medicine*, 11(1), 126. doi:10.1186/1741-7015-11-126

Dalgleish, T. (2004). Cognitive approaches to posttraumatic stress disorder: the evolution of multirepresentational theorizing. *Psychological Bulletin*, 130, 228 -260. doi:

10.1037/0033-2909.130.2.228

de Kleine, R. A., Hendriks, G. J., Kusters, W. J., Broekman, T. G., & van Minnen, A. (2012). A randomized placebo-controlled trial of D-cycloserine to enhance exposure therapy for posttraumatic stress disorder. *Biological Psychiatry, 71*, 962-968.

doi:10.1016/j.biopsych.2012.02.033

de Kleine, R. A., Hendriks, G. J., Smits, J. A., Broekman, T. G., & van Minnen, A. (2014).

Prescriptive variables for d-cycloserine augmentation of exposure therapy for posttraumatic stress disorder. *Journal of Psychiatric Research, 48*, 40-

46. doi:10.1016/j.jpsychires.2013.10.008

de Kleine, R. A., Smits, J. A., Hendriks, G. J., Becker, E. S., & van Minnen, A. (2015).

Extinction learning as a moderator of d-cycloserine efficacy for enhancing exposure therapy in posttraumatic stress disorder. *Journal of Anxiety Disorders, 34*, 63-67. doi:

10.1016/j.janxdis.2015.06.005

Dominique, J. F., Aerni, A., Schelling, G., & Roozendaal, B. (2009). Glucocorticoids and the regulation of memory in health and disease. *Frontiers in Neuroendocrinology, 30*, 358-

370. doi:10.1016/j.yfrne.2009.03.002

Doss, B. D. (2004). Changing the way we study change in psychotherapy. *Clinical Psychology:*

Science and Practice, 11(4), 368–386. doi: 10.1093/clipsy/bph094

DeRubeis, R. J., Gelfand, L. A., German, R. E., Fournier, J. C., & Forand, N. R. (2014).

Understanding processes of change: How some patients reveal more than others—and some groups of therapists less—about what matters in psychotherapy. *Psychotherapy Research, 24*, 419-428. doi: 10.1080/10503307.2013.838654

Difede, J., Cukor, J., Wyka, K., Olden, M., Hoffman, H., Lee, F. S., & Altemus, M. (2014). D-

Difede, J., Cukor, J., Wyka, K., Olden, M., Hoffman, H., Lee, F. S., & Altemus, M. (2014). D-

- cycloserine augmentation of exposure therapy for post-traumatic stress disorder: A pilot randomized clinical trial. *Neuropsychopharmacology*, 39, 1052-1058.
- 10.1038/npp.2013.317
- Ehlers, A., & Clark, D. M. (2000). A cognitive model of posttraumatic stress disorder. *Behaviour Research and Therapy*, 38(4), 319-345. doi: 10.1016/S0005-7967(99)00123-0
- Ehlers, A., Clark, D. M., Hackmann, A., McManus, F., Fennell, M., Herbert, C., & Mayou, R. (2003). A randomized controlled trial of cognitive therapy, a self-help booklet, and repeated assessments as early interventions for posttraumatic stress disorder. *Archives of General Psychiatry*, 60, 1024-1032. doi:10.1001/archpsyc.60.10.1024
- Ehring, T., Kleim, B., & Ehlers, A. (2011). Combining clinical studies and analogue experiments to investigate cognitive mechanisms in posttraumatic stress disorder. *International Journal of Cognitive Therapy*, 4(2), 165-177. doi: 10.1521/ijct.2011.4.2.165
- Foa, E. B., Hembree, E. A., Cahill, S. P., Rauch, S. A., Riggs, D. S., Feeny, N. C., & Yadin, E. (2005). Randomized trial of prolonged exposure for PTSD with and without cognitive restructuring: Outcome at academic and community clinics. *Journal of Consulting and Clinical Psychology*, 73, 953-964. doi: 10.1037/0022-006X.73.5.953
- Foa, E.B., Huppert, J.D., & Cahill, S.P. (2006). Emotional processing theory: An update. In Rothbaum, B.O. (Ed.), *Pathological Anxiety: Emotional processing in etiology and treatment*. (pp. 3-24). New York: Guilford Press.
- Foa, E. B., Ehlers, A., Clark, D. M., Tolin, D. G., & Orsillo, S. M. (1999). The Posttraumatic Cognitions Inventory (PTCI): Development and validation. *Psychological Assessment*, 11, 303–314. doi:10.1037/1040-3590.11.3.303
- Foa, E. B., Hembree, E. A., & Rothbaum, B. O. (2007). *Prolonged exposure therapy for PTSD:*

- Emotional processing of traumatic experiences: Therapist guide.* Treatments that work. New York, NY US: Oxford University Press.
- Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin*, *99*, 20–35. doi: 10.1037/0033-2909.99.1.20
- Foa, E. B., & McLean, C. P. (2015). The efficacy of exposure therapy for anxiety-related disorders and its underlying mechanisms: The case of OCD and PTSD. *Annual Review of Clinical Psychology*, *12*, 1-28. doi: 10.1146/annurev-clinpsy-021815-093533
- Foa, E. B., Molnar, C., & Cashman, L. (1995). Change in rape narratives during exposure therapy for posttraumatic stress disorder. *Journal of Traumatic Stress*, *8*, 675-690. doi: 10.1007/BF02102894
- Foa, E. B., & Rauch, S. A. (2004). Cognitive changes during prolonged exposure versus prolonged exposure plus cognitive restructuring in female assault survivors with posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology*, *72*, 879-884. doi: 10.1037/0022-006X.72.5.879
- Foa, E. B., Riggs, D. S., Massie, E. D., & Yarczower, M. (1995). The impact of fear activation and anger on the efficacy of exposure treatment for posttraumatic stress disorder. *Behavior Therapy*, *26*, 487–499. doi: 10.1016/S0005-7894(05)80096-6
- Galatzer-Levy, I.R., & Bryant, R. A. (2013). 636,120 ways to have posttraumatic stress disorder. *Perspectives in Psychological Science*, *50*, 161-180.
- 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. doi:10.1007/s10608-011-9423-6

- Gallop, R., & Tasca, G. A. (2009). Multilevel modeling of longitudinal data for psychotherapy researchers: II. The complexities. *Psychotherapy Research, 19*(4-5), 438-452.
doi:10.1080/10503300902849475
- Gillihan, S. J., & Foa, E. B. (2011). Fear extinction and emotional processing theory: A critical review. In Schactman, T.R. & Reilly, S.S. (Eds.) *Associative Learning & Conditioning Theory: Human and Non-Human Applications*. New York: Oxford University Press.
- Goldfried, M.R. (2015). On possible consequences of National Institute of Mental Health funding for psychotherapy research and training. *Professional Psychology: Research and Practice, 47*, 77- 83. doi: 10.1037/pro0000034
- Graham, B. M., & Milad, M. R. (2011). The study of fear extinction: Implications for anxiety disorders. *American Journal of Psychiatry, 168*(12), 1255-1265. doi: 10.1176/appi.ajp.2011.11040557
- Guay, S., Marchand, A., & Landry, P. (2010). *Results from a 6-month follow-up of a randomized controlled trial assessing the efficacy of cognitive-behavior therapy combined with D-Cycloserine for treating PTSD*. Presented at the International Society for Traumatic Stress Studies 26th annual meeting, Montreal, Canada, November 4–6, 2010.
- Hagenaars, M. A., Van Minnen, A., & de Rooij, M. (2010). Cognitions in prolonged exposure therapy for posttraumatic stress disorder. *International Journal of Clinical and Health Psychology, 10*(3), 421-434. doi: 10.1037/a0031290
- Halligan, S. L., Michael, T., Clark, D. M., & Ehlers, A. (2003). Posttraumatic stress disorder following assault: The role of cognitive processing, trauma memory, and appraisals. *Journal of Consulting and Clinical Psychology, 71*(3), 419-431. doi: 10.1037/0022-006X.71.3.419

- Harned, M. S., Ruork, A. K., Liu, J., & Tkachuck, M. A. (2015). Emotional activation and habituation during imaginal exposure for PTSD among women with borderline personality disorder. *Journal of Traumatic Stress, 28*(3), 253-257. doi: 10.1002/jts.22013
- Henn-Haase, C., Best, S., Metzler, T. J., Gardner, J., Herbst, E., McCaslin, S., et al. (2010). *Exposure based CBT therapy and D-cycloserine treatment for PTSD in veterans and civilians*. Presented at the International Society for Traumatic Stress Studies 26th annual meeting, Montreal, Canada, November 4–6, 2010.
- Hofmann, S. G. (2008). Cognitive processes during fear acquisition and extinction in animals and humans: Implications for exposure therapy of anxiety disorders. *Clinical Psychology Review, 28*(2), 199-210. doi:10.1016/j.cpr.2007.04.009
- Holtzheimer, P. E., & Mayberg, H. S. (2011). Stuck in a rut: Rethinking depression and its treatment. *Trends in Neurosciences, 34*, 1-9. doi: 10.1016/j.tins.2010.10.004
- Institute of Medicine (IOM). (2008). *Treatment of posttraumatic stress disorder: An assessment of the evidence*. Washington, DC: The National Academies Press.
- Jayawickreme, N., Cahill, S. P., Riggs, D. S., Rauch, S. A., Resick, P. A., Rothbaum, B. O., & Foa, E. B. (2014). Primum non nocere (first do no harm): Symptom worsening and improvement in female assault victims after prolonged exposure for PTSD. *Depression and Anxiety, 31*(5), 412-419. doi: 10.1002/da.22225.
- Jaycox, L. H., Foa, E. B., & Morral, A. R. (1998). Influence of emotional engagement and habituation on exposure therapy for PTSD. *Journal of Consulting and Clinical Psychology, 66*, 185–192. doi: 10.1037/0022-006X.66.1.185
- Jovanovic, T., & Ressler, K. J. (2010). How the neurocircuitry and genetics of fear inhibition may inform our understanding of PTSD. *American Journal of Psychiatry, 167*(6), 648-

- 662.doi: 10.1176/appi.ajp.2009.09071074
- Jun, J. J., Zoellner, L. A., & Feeny, N. C. (2013). Sudden gains in prolonged exposure and sertraline for chronic PTSD. *Depression & Anxiety, 30*, 607-613. doi: 10.1002/da.22119
- Kamprath, K., & Wotjak, C. T. (2004). Nonassociative learning processes determine expression and extinction of conditioned fear in mice. *Learning & Memory, 11*(6), 770-786. doi: 10.1101/lm.86104
- Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology, 3*, 1-27. doi: 10.1146/annurev.clinpsy.3.022806.091432
- Kindt, M. (2014). A behavioural neuroscience perspective on the aetiology and treatment of anxiety disorders. *Behaviour Research and Therapy, 62*, 24-36. doi:10.1016/j.brat.2014.08.012
- Kleim, B., Ehlers, A., & Glucksman, E. (2012). Investigating cognitive pathways to psychopathology: Predicting depression and posttraumatic stress disorder from early responses after assault. *Psychological Trauma: Theory, Research, Practice, and Policy, 4*(5), 527-537. doi:10.1037/a0027006
- Kumpula, M. J., Pentel, K. Z., Foa, E. B., LeBlanc, N. J., Bui, E., McSweeney, L. B., Knowles, K., Bosley, H., Simon, N. M., & Rauch, S. A. M. (2016). Temporal sequencing of change in posttraumatic cognitions and PTSD symptom reduction during prolonged exposure therapy. *Behavior Therapy*. Advanced online publication. doi:10.1016/j.beth.2016.02.008
- Lang, P. J. (1977). Imagery in therapy: An information processing analysis of fear. *Behavior Therapy, 8*, 862-886. doi:10.1016/S0005-7894(77)80157-3
- Lang, P. J. (1979). A bio-informational theory of emotional imagery. *Psychophysiology, 16*(6), 495-512. doi: 10.1111/j.1469-8986.1979.tb01511.x

- LeDoux, J. E. (2014). Coming to terms with fear. *Proceedings of the National Academy of Sciences*, *111*(8), 2871-2878. doi: 10.1073/pnas.1400335111
- Liberzon, I., & Sripada, C. S. (2007). The functional neuroanatomy of PTSD: A critical review. *Progress in Brain Research*, *167*, 151-169. doi: 10.1016/S0079-6123(07)67011-3
- Litz, B. T., Salters-Pedneault, K., Steenkamp, M. M., Hermos, J. A., Bryant, R. A., Otto, M. W., & Hofmann, S. G. (2012). A randomized placebo-controlled trial of D-cycloserine and exposure therapy for posttraumatic stress disorder. *Journal of Psychiatric Research*, *46*(9), 1184-1190. doi:10.1016/j.jpsychires.2012.05.006
- McSweeney, F. K., & Swindell, S. (2002). Common processes may contribute to extinction and habituation. *The Journal of General Psychology*, *129*(4), 364-400. doi: 10.1080/00221300209602103
- Mennin, D., & Farach, F. (2007). Emotion and evolving treatments for adult psychopathology. *Clinical Psychology: Science and Practice*, *14*(4), 329-352. doi: 10.1111/j.1468-2850.2007.00094.x
- McLean, C. P., Su, Y. J., & Foa, E. B. (2015). Mechanisms of symptom reduction in a combined treatment for comorbid posttraumatic stress disorder and alcohol dependence. *Journal of Consulting and Clinical Psychology*, *83*(3), 655-661. doi:10.1037/ccp0000024
- McLean, C. P., Yeh, R., Rosenfield, D., & Foa, E. B. (2015). Changes in negative cognitions mediate PTSD symptom reductions during client-centered therapy and prolonged exposure for adolescents. *Behaviour Research and Therapy*, *68*, 64-69. doi:10.1016/j.brat.2015.03.008
- Michopoulos, V., Norrholm, S. D., & Jovanovic, T. (2015). Diagnostic biomarkers for Posttraumatic Stress Disorder: Promising horizons from translational neuroscience

- research. *Biological Psychiatry*, 78, 344-353. doi:10.1016/j.biopsych.2015.01.005
- Milad, M. R., & Quirk, G. J. (2012). Fear extinction as a model for translational neuroscience: Ten years of progress. *Annual Review of Psychology*, 63, 129-151. doi: 10.1146/annurev.psych.121208.131631
- Moser, J. S., Cahill, S. P., & Foa, E. B. (2010). Evidence for poorer outcome in patients with severe negative trauma-related cognitions receiving prolonged exposure plus cognitive restructuring: Implications for treatment matching in posttraumatic stress disorder. *Journal of Nervous and Mental Disease*, 198(1), 72-75. doi:10.1097/NMD.0b013e3181c81fac
- Murphy, R., Cooper, Z., Hollon, S. D., & Fairburn, C. G. (2009). How do psychological treatments work? Investigating mediators of change. *Behaviour Research and Therapy*, 47(1), 1-5. doi:10.1016/j.brat.2008.10.001
- Myers, K. M., & Davis, M. (2007). Mechanisms of fear extinction. *Molecular Psychiatry*, 12(2), 120-150. doi :10.1038/sj.mp.4001939
- Nacasch, N., Huppert, J. D., Su, Y. J., Kivity, Y., Dinshtein, Y., Yeh, R., & Foa, E. B. (2015). Are 60-minute Prolonged Exposure sessions with 20-minute imaginal exposure to traumatic memories sufficient to successfully treat PTSD? A randomized noninferiority clinical trial. *Behavior Therapy*, 46(3), 328-341. doi:10.1016/j.beth.2014.12.002
- O'Kearney, R., & Perrott, K. (2006). Trauma narratives in posttraumatic stress disorder: A review. *Journal of Traumatic Stress*, 19(1), 81-93. doi: 10.1002/jts.20099
- Øktedalen, T., Hoffart, A., & Langkaas, T. F. (2015). Trauma-related shame and guilt as time-varying predictors of posttraumatic stress disorder symptoms during imagery exposure and imagery rescripting—A randomized controlled trial. *Psychotherapy Research*, 25, 1-

15. doi: 10.1080/10503307.2014.917217
- Pincus, H.A., & England, M.J. (2015). Improving the quality of psychosocial interventions for mental and substance use disorders: A report from the IOM. *Journal of the American Medical Association, 314*, 1227-1228. doi: 10.1001/jama.2015.9212
- Pitman, R. K., Orr, S. P., Altman, B., Longpre, R. E., Poiré, R. E., Macklin, M. L., ... & Steketee, G. S. (1996). Emotional processing and outcome of imaginal flooding therapy in Vietnam veterans with chronic posttraumatic stress disorder. *Comprehensive Psychiatry, 37*(6), 409-418. doi: 10.1016/S0010-440X(96)90024-3
- Power, M., & Dalgleish, T. (2015). *Cognition and emotion: From order to disorder*. 3rd edition. Psychology Press: New York, NY.
- Powers, M. B., Halpern, J. M., Ferenschak, M. P., Gillihan, S. J., & Foa, E. B. (2010). A meta-analytic review of prolonged exposure for posttraumatic stress disorder. *Clinical Psychology Review, 30*(6), 635-641. doi: 10.1016/j.cpr.2010.04.007
- Rauch, S. M., Foa, E. B., Furr, J. M., & Filip, J. C. (2004). Imagery vividness and perceived anxious arousal in prolonged exposure treatment for PTSD. *Journal of Traumatic Stress, 17*, 461-465. doi:10.1007/s10960-004-5794-8
- Reist, C., Gory, A., & Hollifield, M. (2017). Sleep-disordered breathing impact on efficacy of prolonged exposure therapy for posttraumatic stress disorder. *Journal of Traumatic Stress*. Advanced online publication. doi: 10.1002/jts.22168
- Resick, P. A., & Schnicke, M. K. (1992). Cognitive processing therapy for sexual assault victims. *Journal of Consulting and Clinical Psychology, 60*(5), 748-756.
doi:10.1037/0022-006X.60.5.748
- Romano, M. & Peters, L. (2015). Evaluating the mechanisms of change in motivational

- interviewing in the treatment of mental health problems: A review and meta-analysis. *Clinical Psychology Review*, 38, 1-12. doi: 10.1016/j.cpr.2015.02.008
- Rothbaum, B. O., Price, M., Jovanovic, T., Norrholm, S. D., Gerardi, M., Dunlop, B., ... & Ressler, K. J. (2014). A randomized, double-blind evaluation of D-cycloserine or alprazolam combined with virtual reality exposure therapy for posttraumatic stress disorder in Iraq and Afghanistan War veterans. *American Journal of Psychiatry*, 171, 640 - 648. doi: 10.1176/appi.ajp.2014.13121625
- Schottenbauer, M. A., Glass, C. R., Arnkoff, D. B., Tendick, V., & Gray, S. H. (2008). Nonresponse and dropout rates in outcome studies on PTSD: Review and methodological considerations. *Psychiatry: Interpersonal and Biological Processes*, 71(2), 134-168. doi: 10.1521/psyc.2008.71.2.134
- Smits, J. A., Julian, K., Rosenfield, D., & Powers, M. B. (2012). Threat reappraisal as a mediator of symptom change in cognitive-behavioral treatment of anxiety disorders: A systematic review. *Journal of Consulting and Clinical Psychology*, 80(4), 624-635. doi:10.1037/a0028957
- Smits, J. A. J., Rosenfield, D., Otto, W. M., Powers, M. B., Hofmann, S. G., Telch, M. J., Pollack, M. H., & Tart, C. D. (2013). D-Cycloserine enhancement of fear extinction is specific to successful exposure sessions: Evidence from the treatment of height phobia. *Biological Psychiatry*, 73(11), 1054-1058. doi:10.1016/j.biopsych.2012.12.009
- Sripada, R. K., & Rauch, S. A. (2015). Between-session and within-session habituation in Prolonged Exposure Therapy for posttraumatic stress disorder: A hierarchical linear modeling approach. *Journal of Anxiety Disorders*, 30, 81-87. doi:10.1016/j.janxdis.2015.01.002

- Tart, C. D., Handelsman, P. R., DeBoer, L. B., Rosenfield, D., Pollack, M. H., Hofmann, S. G., ... & Smits, J. A. (2013). Augmentation of exposure therapy with post-session administration of D-cycloserine. *Journal of Psychiatric Research, 47*(2), 168-174. doi:10.1016/j.jpsychires.2012.09.024
- Tronson, N. C., & Taylor, J. R. (2007). Molecular mechanisms of memory reconsolidation. *Nature Reviews Neuroscience, 8*(4), 262-275. doi:10.1038/nrn2090
- van der Velden, A. M., Kuyken, W., Wattar, U., Crane, C., Pallesen, K. J., Dahlgard, J., Fjorback, L. O., & Piet, J. A systematic review of mechanisms of change in mindfulness-based cognitive therapy in the treatment of recurrent major depressive disorder. *Clinical Psychology Review, 37*, 26-39. doi:10.1016/j.cpr.2015.02.001
- van Minnen, A., Arntz, A., & Keijsers, G. P. J. (2002). Prolonged exposure in patients with chronic PTSD: Predictors of treatment outcome and dropout. *Behaviour Research and Therapy, 40*, 439-457. doi: 10.1016/S0005-7967(01)00024-9
- van Minnen, A. & Foa, E. B. (2006). The effect of imaginal exposure length on outcome of treatment for PTSD. *Journal of Traumatic Stress, 19*(4), 427-438. doi: 10.1002/jts.20146
- van Minnen, A., & Hagenaars, M. (2002). Fear activation and habituation patterns as early process predictors of response to prolonged exposure treatment in PTSD. *Journal of Traumatic Stress, 15*(5), 359-367. doi: 10.1023/A:1020177023209
- van Minnen, A., Hendriks, L., & Olf, M. (2010). When do trauma experts choose exposure therapy for PTSD patients? A controlled study of therapist and patient factors. *Behaviour Research and Therapy, 48*(4), 312-320. doi:10.1016/j.brat.2009.12.003
- van Minnen, A., Wessel, I., Dijkstra, T., & Roelofs, K. (2002). Changes in PTSD patients' narratives during prolonged exposure therapy: A replication and extension. *Journal of*

- Traumatic Stress*, 15(3), 255-258. doi:10.1023/A:1015263513654
- Walz, L. C., Nauta, M. H., & Aan Het Rot, M. (2014). Experience sampling and ecological momentary assessment for studying the daily lives of patients with anxiety disorders: A systematic review. *Journal of Anxiety Disorders*, 28(8), 925-937.
doi:10.1016/j.janxdis.2014.09.022
- Wisco, B. E., Sloan, D. M., & Marx, B. P. (2013). Cognitive emotion regulation and written exposure therapy for posttraumatic stress disorder. *Clinical Psychological Science*, 1(4), 435–442. doi:10.1177/2167702613486630
- Wolpe, J., & Lazarus, A.A. (1966). *Behavior therapy techniques*. New York: Pergamon Press.
- Yehuda, R., Daskalakis, N. P., Desarnaud, F., Makotkine, I., Lehrner, A. L., Koch, E., ... & Bierer, L. M. (2013). Epigenetic biomarkers as predictors and correlates of symptom improvement following psychotherapy in combat veterans with PTSD. *Frontiers in Psychiatry*, 4, 118. doi: 10.3389/fpsy.2013.00118
- Yehuda, R., Bierer, L. M., Pratchett, L. C., Lehrner, A., Koch, E. C., Van Manen, J. A., ... & Hildebrandt, T. (2015). Cortisol augmentation of a psychological treatment for warfighters with posttraumatic stress disorder: Randomized trial showing improved treatment retention and outcome. *Psychoneuroendocrinology*, 51, 589-597. doi: 10.1016/j.psyneun.2014.08.004
- Zalta, A. K. (2015). Psychological mechanisms of effective cognitive–behavioral treatments for PTSD. *Current Psychiatry Reports*, 17(4), 1-8. doi: 10.1007/s11920-015-0560-6
- Zalta, A. K., Gillihan, S. J., Fisher, A. J., Mintz, J., McLean, C. P., Yehuda, R., & Foa, E. B. (2014). Change in negative cognitions associated with PTSD predicts symptom reduction in prolonged exposure. *Journal of Consulting and Clinical Psychology*, 82(1), 171-175.

doi: 10.1037/a0034735

Zoellner, L. A., & Foa, E. B. (2016). Applying Research Domain Criteria (RDoC) to the study of fear and anxiety: A critical comment. *Psychophysiology*, *53*(3), 332-335. doi: 10.1111/psyp.12588

Table 1. Kazdin's (2007) Seven Requirements for Demonstrating Mediators and Mechanisms of Change

Criterion	Concept
<i>Strong association</i>	Demonstrates a strong association (correlation) between the intervention (A) and hypothesized mediator (B). Ideally, also demonstrates a strong association between the proposed mediator (B) and the outcome (C) (e.g., symptom reduction).
<i>Specificity</i>	Demonstrates a specific effect, whereby the proposed mediator is shown to account for therapeutic change to a greater degree than other plausible constructs when effects are compared.
<i>Consistency</i>	Demonstrates evidence of a consistent relationship between mediator and outcome, by way of replication across studies, samples and treatment conditions. Does not rule out possibility of moderation to explain between-study differences.
<i>Experimental manipulation</i>	Uses experimental design involving either randomization to treatments (e.g., randomized controlled trials) to demonstrate connection between intervention (A) and outcome (C), or (less commonly) experimental manipulation of the proposed mediator (B) in relation to outcomes (C).
<i>Timeline (Temporal precedence)</i>	Demonstrates a plausible causal or mediating relationship on the basis of timing of measurements. That is, causal forces and mediators must temporally precede the effects and outcomes they are expected to act upon.
<i>Gradient</i>	Demonstrates evidence of a graded relationship, whereby stronger “doses” of a proposed mediator are associated with a greater change in the outcome.
<i>Plausibility / Coherence</i>	Offers an explanatory model that integrates with broader scientific knowledge base, and regarded as reasonable and coherent with other relevant evidence.

Content adapted from Kazdin (2007).

Table 2. Studies Evaluating Potential Mechanisms of Exposure Therapy for PTSD

Study	Population (N)	Treatments (Context)	Outcome / Model	Variable	Measure / Test	Finding	QE	SE
Bedard-Gilligan et al. (in press)	Mixed trauma (77)	PE vs sertraline (RCT subsample)	Mixed effects regression of PSS-SR	OTN	Characteristics of narratives (trauma, positive, negative)	Treatment (PE vs sertraline) or response status did not predict reduced fragmentation of post-treatment trauma narratives	--	--
Bluett et al. (2014)	Mixed trauma (88)	PE (RCT subsample)	Pre-post residualized PSS-I; loss of PTSD diagnosis at post	BSH	Reliable change in SUDS (peak, mean) from 1st IE to final IE	Patients with SUDS reliable change had greater PTSD change; no differences in % retaining diagnosis	4	2
Cooper et al. (in press)	Mixed trauma (134)	PE vs sertraline (RCT subsample)	Lagged multilevel model of residualized PSS-SR	TRBC	PTCI ratings at each session	Belief change predicted next-session PTSD change, more in PE vs sertraline	8	3
deKleine et al.(2015)	Mixed trauma (67)	PE+ DCS vs PE+PBO (RCT)	Pre-post CAPS change; response, remission (<20 at post)	FIL	DCS vs PBO treatment effect ¹	No differences in CAPS improvement or remission; DCS group more likely to respond	6	0
				FIL	DCS vs PBO: end SUDS ratings at each session	End SUDS predicted greater session-level and post-treatment improvement; no differences between DCS and PBO	11	0
				BSH/FIL ²	DCS vs PBO: mean change in peak SUDS between sessions	BSH predicted greater post-treatment improvement; no differences between DCS and PBO	8	2
				WSH/FIL ²	DCS vs PBO: peak-end SUDS per session, mean across treatment	WSH predicted greater session-level and post-treatment improvement; no differences between DCS and PBO	8	1
Difede et al. (2013)	9/11 WTC-related trauma (25)	VRE+DCS vs VRE+PBO (RCT)	Pre-post and pre-6 month FU CAPS; remission (<20 at post)	FIL	DCS vs PBO treatment effect	Groups not different on CAPS at post, DCS superior at FU; higher remission in DCS at post and FU	3	0
Foa, Molnar & Cashman (1995)	Women with SA-related trauma (14)	PE (open trial)	Pre-post change in composite trauma-related anxiety (%)	OTN	Change in fragmentation & organization, 1 st to last narrative	Change in fragmentation correlated with reductions in trauma-related anxiety. Change in organization not significantly related.	--	--
Foa & Rauch (2004)	Women with SA-related trauma (54)	PE vs PE+CR (RCT, pooled sample)	Pre-post residualized PSS-I	TRBC	Pre-post residualized PTCI	Belief change related to PTSD change	5	2
Foa, Riggs et al. (1995)	Women with SA-related trauma (12)	PE (open trial)	Pre-post percent improvement in PSS	EE	(A) observer rated FF matched to (B) peak SUDS, 1st IE	PSS improvement correlated with and peak SUDS	2	2
Gallagher & Resick (2012)	Women with SA-related trauma(88)	PE (RCT subsample)	Pre-post change in CAPS	BSH	Mean SUDS, 1st IE - final IE	BSH correlated with PTSD change	5	1

Hagenaars et al. (2010)	Mixed trauma (77)	PE (open trial)	Pre-post, pre-6 month FU residualized PSS-I	TRBC	Residualized PTCI: pre-post and pre-6 month FU	Belief change related to PTSD change: pre-post and pre-FU	3	2
Harned et al (2015)	Women with comorbid BPD (16)	PE (open trial)	Loss of PTSD diagnosis at post	EE	Pre-to-peak SUDS, averaged across all IE; highest emotion rating for fear, sadness, anger, guilt, shame, disgust	No measures of EE predicted loss of PTSD diagnosis at post.	1	0
				WSH	Peak-to-post SUDS, averaged across all IE; pre-to-post change in specific emotions	No measures of WSH predicted loss of PTSD diagnosis at post.	1	0
				BSH	Peak SUDS, 1st IE - final IE, averaged across different trauma memories; similar test for specific emotions	Patients who retained their PTSD diagnosis had less BSH for SUDS, sadness, and anger	1	2
Jaycox et al. (1998)	Women with SA-related trauma(37)	PE (open trial)	Post-treatment GESF	BSH	Cluster analysis of mean peak-end SUDS per IE, across all sessions	Greater proportion meeting GESF in High SUDS/Declining group vs High SUDS/Stable or Low SUDS/Stable groups	2	0
				WSH	Peak - end SUDS (mean across treatment)	Odds of GESF unrelated to WSH, IE duration		
Kumpula et al. (2016)	Mixed trauma (46)	PE (RCT subsample)	Lagged multilevel model of residualized PDS	TRBC	PTCI ratings at baseline, sessions 2,4,6, 8, with post-hoc subscale analyses	PTCI changed preceded PTSD change, with strongest effects for negative beliefs about self; no change in self-blame	6	1
Litz et al. (2012)	Male Veterans with mixed trauma (26)	4 session IE+DCS vs IE+PBO (RCT)	Post-treatment CAPS, treatment response (10+ reduction from pre)	FIL	DCS vs PBO treatment effect	Greater improvements in PBO vs DCS, with higher response (70% vs 30%)	4	-1 ³
McLean et al. (2015)	Mixed trauma with comorbid AD (159)	PE+NAL, PE+PBO, SC+NAL, SC+PBO (RCT)	Lagged multilevel, mediation analyses of PSS-I	TRBC	PTCI ratings every 4 weeks	Belief change mediated PTSD improvements in PE, with evidence of bidirectional effects	7	2
Nacasch et al. (2015)	Veterans with mixed trauma (39)	PE 60min vs 90min (RCT)	Pre-post residualized PSS-I	TRBC	Pre-post residualized PTCI	Belief change related to PTSD change	9	2
				BSH	Peak SUDS, 1st IE - final IE	PSS-I change significantly correlated with BSH; BSH greater in 90-min vs 60-min	10	2
				WSH	Peak - end SUDS, 1st IE and final IE	PSS-I change not correlated with WSH at 1st IE or final IE; WSH greater in 90-min vs 60-min at 1st IE	9	1

Øktedalen et al (2015)	Inpatients with mixed trauma (65)	PE vs IR (RCT)	Lagged multilevel model of residualized PSS-SR	TRBC	PTCI subset assessed weekly, modeled as within- and between-patient differences	Within-patient belief change predicts subsequent PTSD change in both PE and IR	8	1
Pitman et al (1996)	Male Veterans (20)	Flooding therapy (open trial)	Pre-post change in avoidance and intrusive symptoms	BSH	HR increase (pre to peak), 1st flooding - last flooding	BSH predicts fewer intrusions	0	1
				WSH	HR decrease (peak - end of flooding)	WSH correlated with less intrusions; p<.05 but noted as non-significant	0	2
				EE	HR increase (imaginal peak - resting baseline)	HR increase correlated with fewer intrusions	0	2
Rauch et al. (2004)	Women with SA- or IPV-related trauma (69)	PE vs PE+CR (RCT, pooled sample)	Pre-post residualized PSS-I (or post PSS-I)	BSH	Peak SUDS, 1st IE - final IE	Greater BSH correlated with PTSD change	3	2
				EE	Peak SUDS, 1st & final IE	Post PSS-I correlated with peak SUDS at final, not 1st session	3	1
Rothbaum et al. (2014)	Veterans with mixed trauma (156)	5 sessions VRE+DCS vs VRE+PBO vs VRE+ALP (RCT)	CAPS, PSS-SR at post, 12 month FU	FIL	DCS vs PBO treatment effect; relationship to cortisol & startle	No difference on CAPS or PSS-SR, at post or FU; DCS has better cortisol response, attenuated startle	5	0
				FIL	DCS vs other treatments: Mean between-session change in peak SUDS across treatment	BSH predicts CAPS at post but only in DCS condition	6	3
Sripada & Rauch (2015)	Veterans with mixed trauma (12)	PE (RCT subsample)	Pre-post change in CAPS; responder status (50% reduction)	BSH	All SUDS data from IEs; HLM model of between-session SUDS slope	CAPS change, responder status significantly associated with between-session SUDS slope	6	3
				WSH	HLM of all IE SUDS, within-session slope	CAPS change and responder status not significantly associated with slope of within-session SUDS change	6	0
van Minnen & Foa (2006)	Mixed trauma (92)	PE 60min vs 90min (non-randomized open comparison)	PSS-SR at final session, 1 month FU	BSH	Peak SUDS, 1st IE - final IE	BSH correlated with PTSD symptoms at post and FU; BSH did not differ by 60- or 90-min sessions	4	2
				WSH	Peak - end SUDS, 1st IE and final IE	WSH (1 st , final) unrelated to PSS-SR (post, FU); more WSH in 90-min vs 60-min protocol	4	0

van Minnen & Hagens (2002)	Mixed trauma (34)	PE (open trial)	Improved/not improved per post-treatment GESF; pre-post residualized PSS-SR	EE	(A) Peak - pre SUDS, 1st IE; (B) Peak - pre SUDS at 2nd IE; (C) Mean peak - mean pre SUDS, IE homework	No differences between improved / not improved after controlling for pre-IE SUDS	4	0
				WSH	(A) Peak - end SUDS, 1st IE, 2nd IE; (B) Mean peak - mean end SUDS, IE homework	Improved patients had greater WSH for homework, but WSH not correlated with residualized PSS-I change	4	0
				BSH	(A) Peak SUDS 1st IE - Peak SUDS 2nd IE, (B) Peak SUDS 1st IE - Mean Peak SUDS Homework	Improved patients higher on all measures of BSH. PSS-SR change correlated with BSH	3	1
van Minnen, et al. (2002)	Mixed trauma (20)	PE (open trial)	Improved/not improved per post-treatment GESF	OTN	Change from 1 st to last narrative in fragmentation, organization, other indices	No difference between groups on organization, fragmentation; improved group has greater reduction in disorganized thoughts	--	--
Yehuda et al. (2015)	Veterans with mixed trauma (24)	PE+HC vs PE+PBO (RCT)	CAPS, PSS-SR at post, 12 month FU	FIL	HC vs PBO treatment effect	No differences in CAPS; HC had greater retention, superior outcomes for completers	5	0
Zalta et al. (2014)	Women with mixed trauma (64)	PE (RCT subsample)	Lagged multilevel model of residualized PSS-SR	TRBC	PTCI ratings at each session	Belief change predicted next-session PTSD change, supporting expected temporal relation	7	1

Note: Study *N* based on primary analyses of mechanism, using largest value if more than one subset. For SE and QE ratings, primary tests and designation of mechanisms reported in Online Supplement. The largest reported effect size was used if multiple comparisons were conducted without specifying primary test.

Rating Variables: SE = strength of evidence (range: 0-3); QE = quality of evaluation (range 0-14). **Mechanism Variables:** EE = emotional engagement; FIL = fear inhibition learning; BSH = between-session habituation; OTN = organization of trauma memory; TRBC = trauma-related belief change; WSH = within-session habituation. **Symptom Measures:** CAPS = Clinician Assessed PTSD Scale; PDS = Posttraumatic Stress Diagnostic Scale; PSS = Posttraumatic Symptom Scale, (-SR = Self-Report, -I = Interview); PTCI = Posttraumatic Cognitions Inventory. **Study Variables:** AD = alcohol dependence; BPD = borderline personality disorder; FF = facial fear; FU = follow-up phase; GESF = good end state functioning; HR = heart rate; IE = imaginal exposure; IPV = interpersonal violence; RCT = randomized clinical trial; SA = sexual assault; SUDS = subjective units of distress scale. **Treatment Variables:** ALP = alprazolam; CR = cognitive restructuring; DCS = d-cycloserine; HC = hydrocortisone; IR = imagery rescripting; NAL = naltrexone; PE = prolonged exposure; SC = supportive counseling; VRE = virtual reality exposure

¹ As reported in deKleine et al (2012)

² SE and QE scores reported here for BSH and WSH comparisons (collapsed across groups). See Online Supplement.

³ Negative score reflects significant finding opposite to predicted direction. See Online Supplement.

Table 3. Recommendations for Future Mechanism Research on Prolonged Exposure

Issue	Field-Level Changes	Study-Level Changes
<i>Results cannot easily be compared across studies due to substantial method variance.</i>	Develop standard definitions of (a) mechanism variables and (b) outcomes, including guidelines for timing of assessments	Utilize standard methods of assessment, or include a comparison to these if piloting a novel approach to measuring mechanism and/or outcomes
<i>Mechanisms are assessed using a single response modality.</i>	Develop pragmatic, translational paradigms to investigate understudied mechanisms in subclinical and analogue samples	Use a complementary alternative method to assess key mechanisms(e.g., observer rating, challenge paradigm)
<i>Studies use low-tech analyses that may obscure or distort mechanism-outcome relationships</i>	Develop guiding document for analytic frameworks that are (a) robust to missing data, (b) consider temporal relationships, and (c) facilitate comparative designs	Assess mechanism and outcome variables at multiple time points, ideally also considering confounding factors
<i>Studies rarely address intra-individual differences (stable or treatment related)</i>	Promote research on variables that predict (a) overall propensity to respond and (b) specific mechanism-outcome relationships	Evaluate the impact of implementation differences (e.g., exposure duration) & mechanism-specific confounds (e.g., anticipatory anxiety)
<i>Studies offer inadequate tests of mechanistic relationships</i>	Develop quality rating system based on key components of adequate mechanism research (e.g., Kazdin, 2007)	Explicitly describe how analyses will address questions of mechanism (e.g., which pathways & processes)
<i>Studies are often underpowered to detect mechanism effects</i>	Encourage and facilitate open access to large datasets to enhance power to detect moderate-sized effects	Require <i>a priori</i> power analyses for studies of mechanistic relationships, including handling of attrition

Role of funding sources

This work was supported by the National Institute of Mental Health (R01 MH066348) awarded to NCF. The funding source had no role in study design, analysis, writing or decision to submit this paper for publication.

Contributors

AAC and EGC developed the concept for the study and reviewed the literature. NCF consulted on the project concept and framework for the review. All three authors contributed to writing the manuscript, and approve of its final version.

Conflict of interest

All authors declare that they have no conflicts of interest.

Acknowledgments

The authors wish to thank members of the PTSD Treatment and Research Program for their helpful feedback on this project.

Highlights

- Empirical evidence for six putative PE mechanisms is reviewed
- Belief change and between-session habituation have strongest evidence base
- Extinction and emotional engagement have an intermediate level of evidence
- Trauma narrative change and within-session habituation have weak evidence base
- Recommendations for future mechanism studies are discussed