Review

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

Andrew A. Cooper[a,b,*], Erin G. Clifton[b], Norah C. Feeny[b]

[a] Department of Psychology, University of Toronto Scarborough, Canada
[b] Department of Psychological Sciences, Case Western Reserve University, Cleveland, Ohio, USA

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.

ARTICLE INFO

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.

1. 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

Table 1
Kazdin's (2007) seven requirements for demonstrating mediators and mechanisms of change.
Content adapted from Kazdin (2007).

Criterion	Concept
Strong association	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).
Specificity	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.
Consistency	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.
Experimental manipulation	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).
Timeline (temporal precedence)	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 upon which they are expected to act.
Gradient	Demonstrates evidence of a graded relationship, whereby stronger "doses" of a proposed mediator are associated with a greater change in the outcome.
Plausibility/coherence	Offers an explanatory model that integrates with broader scientific knowledge base, and regarded as reasonable and coherent with other relevant evidence.

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 reducing 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) proposed seven explicit criteria for evaluating mediators as part of a framework for investigating mechanisms (see Table 1). Unfortunately, there are a myriad of conceptual[1] 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 topic, 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, 1999), 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 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 (2007) seven key criteria (see 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.

2. Description of prolonged exposure therapy

PE is a manualized cognitive-behavioral intervention for PTSD (e.g., Foa et al., 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 train, 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

---

[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.