Optimizing Exposure Therapy for Anxiety Disorders: An Inhibitory Learning and Inhibitory Regulation Approach

Michelle Craske
University of California, Los Angeles, CA, USA

Keywords
Exposure · Extinction · Inhibition · Anxiety

Summary
Exposure-based cognitive-behavioral therapy is the treatment of choice for anxiety disorders. Unfortunately, many individuals fail to benefit from treatment or demonstrate a return of fear. Inhibitory learning and inhibitory regulation models provide a parsimonious and unifying framework from which to situate exposure therapy and provide useful strategies to augment exposure therapy, especially as individuals with anxiety disorders show inhibitory deficits. This paper provides an overview of our application of inhibitory learning from the science of extinction, and inhibitory regulation from the neuroscience of emotion regulation, to exposure therapy for anxiety disorders.

Introduction
Exposure therapy refers to repeated, systematic exposure to cues that are feared, avoided, or endured with dread. Exposure includes in-vivo exposure to actual objects, events or situations, imaginal exposure to traumatic memories, catastrophic images or obsessional thoughts; and interoceptive exposure to bodily sensations. Several meta-analyses have demonstrated that exposure therapy is highly effective, whether alone or in combination with cognitive restructuring or somatic strategies [Norton and Price, 2007; Hofmann and Smits, 2008]. However, exposure therapy is not effective for everyone. For example, a number of individuals refuse to begin exposure therapy, with estimates ranging from 25% [Garcia-Palacios et al., 2007] to 30% [Issakidis and Andrews, 2004]. Of those who enter exposure-based cognitive behavior therapy, a number fail to complete treatment, with estimates ranging from 15 to 30% [e.g., Haby et al., 2006]. In terms of treatment success, ap-
proximately 50% fail to achieve clinically significant improvement with cognitive behavioral therapy for anxiety disorders, with higher rates in intent-to-treat samples (Loerinc et al., under review). Finally, a portion of individuals (19–62%) demonstrate a return of fear to particular targets of exposure therapy [Craske and Mystkowski, 2006].

Mechanisms of Exposure Therapy: Inhibitory Model of Extinction

What are the mechanisms of exposure therapy and how can those mechanisms be optimized in such a way that enhance the response rate and reduce the return of fear? Habitation, or reduction in response strength with repeated stimulus presentations [Groves and Thompson, 1970], provides a descriptive framework for responses during exposure therapy because self-reported fear and physiological arousal most often decline within and across exposure occasions [e.g., Kozak et al., 1988]. The very influential ‘emotional processing theory’ [Foote and Kozak, 1986; Foote and McNally, 1996] emphasized within-session habituation (i.e. fear reduction from initial activation or peak fear to end of exposure) as a precursor to between-session habituation (i.e. fear reduction from one exposure session to the next) and long-term cognitive correction. In accord, clinicians are typically trained to use initial elevation of fear followed by within- and between-session reductions in fear as signs of treatment success, and to continue an exposure trial until fear declines by at least 50% of initial levels. Although enticing in its face validity, support for the theory has been inconsistent at best [Craske et al., 2008a]. Rather, the evidence suggests that the amount by which fear declines from the beginning to end of an exposure session (i.e. within-session habituation) does not significantly predict self-report questionnaires or behavioral avoidance testing at follow-up, covarying baseline levels on those measures [Craske et al., 2008a; Baker et al., 2010; Culver et al., 2012; Kircanski et al., 2012].

If habituation is not the critical index of learning, then what is? A return to the science of fear learning and extinction may offer answers. It is now thought that inhibitory learning is central to extinction [Bouton, 1993; Miller and Matzel, 1988; Wagner, 1981], although additional mechanisms, such as habituation, may be at play as well [Myers and Davis, 2007]. Within a Pavlovian conditioning approach, the inhibitory learning models mean that the original CS-US (conditioned stimulus-unconditioned stimulus) association learned during threat conditioning is not erased during extinction, but rather is left intact as a new, secondary inhibitory learning about the CS-US develops [e.g., Bouton and King, 1983; Bouton, 1993]. In other words, after extinction, the CS possesses 2 meanings: its original excitatory meaning (CS-US) as well as an additional inhibitory meaning (CS-noUS). Therefore, even though fear subsides with enough trials of the CS in the absence of the US (i.e. extinction), retention of at least part of the original association can be uncovered by various procedures, with each one showing a continuing effect of the original excitatory association. First, conditional fear shows spontaneous recovery [Quirk, 2002], meaning that the strength of the CR increases in proportion to the amount of time between the end of extinction training and retest. Second, renewal of conditional fear occurs if the surrounding context is changed between extinction and retest [Bouton, 1993]. In other words, extinction effects are specific to the context in which extinction occurs [e.g., Vansteenwegen et al., 2005; Alvarez et al., 2008; Milad et al., 2007; Neumann and Longbottom, 2008]. Such context specificity has been observed in clinical analog samples undergoing exposure therapy and follow-up testing in the same versus different contexts [Mystkowski et al., 2002; 2003; 2006; Culver et al., 2011]. Third, reinstatement of conditional fear occurs if unsigned (or unpaired) US presentations occur in between extinction and retest [Rescorla and Heth, 1975], although the effect depends on the fearfulness of the context in which the CS is tested. Reinstatement is well established in animal threat conditioning studies and more recently has been shown in human conditioning studies [e.g., Hermans et al., 2005; Van Damme et al., 2006; Zbozinek et al., 2014], especially when the CS retains negative valence following extinction. Fourth, rapid reacquisition of the CR is seen if the US-pairings are repeated following extinction [Ricker and Bouton, 1996; Culver et al., 2014]; being more rapid than the original learning indicates the carry-over effects of the original acquisition learning.

Notably, a number of threat conditioning and extinction studies have shown that the level of fear (or freezing in rodents) at the end of extinction training is not predictive of fear/freezing at retest [e.g., Rescorla, 2006; Plendl and Wotjak, 2010; Prenoveau et al., 2013]. In fact, ‘extinction failure’ versus ‘extinction success’ groups of rodents, defined by responding at retest, showed identical patterns of extinction training freezing [Peters et al., 2010]. Together, these findings from threat-conditioning studies in animals and humans confirm our findings in clinical samples reported earlier.

Figure 1 provides a clinical translation of the inhibitory extinction model as it applies to exposure therapy. An individual enters treatment with a threat expectancy, such as ‘If I panic, I might die’ or ‘If I am socially rejected, it would be unbearable’. As a result of exposure therapy, a competing non-threat expectancy develops, such as ‘If I panic, I am unlikely to die’ or ‘If I am socially rejected, I will survive’. After completion of exposure therapy, the level of fear that is experienced when the stimulus is re-encountered is dependent upon which expectancy is activated. Activation of the original threat expectancy will enhance the expression of fear whereas activation of the exposure-based non-threat expectancy will lessen fear expression. Which of those 2 expectancies are more strongly activated will depend on the factors outlined earlier. For example, lengthy intervals of time since the completion of exposure therapy will give preference to the original threat expectancy, in line with spontaneous recovery. Thus, the patient whose fear of air travel significantly reduces by the end of treatment is likely to report a return in their fear of flying if they do not continue to practice air travel once treatment is completed. Or, if exposure therapy is completed in one or only a limited number of contexts (such as in the presence of a therapist, or always immediately preceding or following a therapy session), the original threat expectancy memory is
likely to be activated when the phobic stimulus is subsequently encountered in a different context (such as when alone or when unrelated to a therapy session). Further, unpredicted adverse events following exposure therapy can reactivate the original threat expectancy for a given stimulus, depending on context. For example, fear of asking questions in work meetings may resurface after being criticized unexpectedly in another social situation, or possibly even after an unrelated adverse event such as a motor vehicle accident [Sokol and Lovibond, 2012]. Finally, the original threat expectancy is more likely to be reactivated with re-traumatization, as may occur in combat situations or other dangerous environments.

Notably, the level of fear expressed during or at the end of exposure therapy has no relationship to the factors responsible for which expectancy is activated (and therefore level of fear expressed) at retest. Such a discord between level of fear expressed and underlying associative mechanisms has been well established [LeDoux, 2014].

Furthermore, the brain-behavior associations underlying fear conditioning and extinction have been well mapped. The amygdala plays a primary role in fear conditioning, along with the hippocampus, which is involved in processing contextual cues of conditioning, the insular cortex, which is involved in interoception and awareness of and sensitivity to visceral activity and the dorsal and rostral ACC, which appears to have a role in anticipation of the CS and US [Shin and Liberzon, 2010]. The ventromedial prefrontal cortex (vmPFC) is believed to mediate extinction, and the hippocampus to modulate extinction by providing information regarding safe versus dangerous contexts [e.g., Kalisch et al., 2006; Milad et al., 2007; Milad et al., 2009]. It is theorized that the PFC exerts inhibitory control over the amygdala through vmPFC activation of inhibitory lateral nucleus interneurons or inhibitory projections to the central nucleus of the amygdala. That is, the PFC serves as the neurobiological basis for inhibitory learning. Also, it is presumed that the hippocampus creates a unique representation of the context in which extinction took place, and modulates the effect of the PFC on the amygdala [Quirk and Mueller, 2008]. In sum, the neurobiology of extinction supports the inhibitory learning believed to take place during extinction.

**Deficits in Inhibition in Anxiety Disorders**

Individuals with anxiety disorders show deficits in extinction learning and related mechanisms, the very mechanisms that we are attempting to target through exposure therapy. A meta-analysis of behavioral studies showed that individuals with a variety of anxiety disorders exhibit stronger responding to the CS+ not only during conditioning but also during extinction relative to controls [Lissek et al., 2005]. Furthermore, in differential conditioning paradigms, they exhibit higher levels of responding to the CS− (the stimulus never paired with the US), during conditioning and extinction, relative to controls. Since the meta-analysis, additional studies have confirmed either weakened extinction (i.e. elevated responding to the CS+ during extinction training), long term deficits in extinction (i.e., elevated responding to the CS+ at extinction retest), or elevated responding to the CS−, in anxiety disorders [Blechert et al., 2007; Craske et al., 2008b; Waters et al., 2009; Lissek et al., 2009; Michael et al., 2007; Milad et al., 2009].

One explanation of these findings is that individuals with anxiety disorders show impaired inhibitory learning [Lissek et al., 2005]. Indeed, impaired inhibitory (or safety) learning to either a stimulus that does not signal threat (CS−) or to a stimulus that no longer signals threat (CS+ during extinction) has been theorized to be central to the pathology of anxiety disorders [Davis et al., 2000; Craske et al., 2009a]. Furthermore, in children at risk for anxiety disorders these difficulties are present, suggesting that they may serve as a pathway for the acquisition of anxiety disorders [Craske et al., 2008b]. In parallel, we have observed increases in defensive responding to safe cues within a fear-potentiated startle paradigm (akin to elevated responding to the CS− in differential extinction) as a correlate of the trait of neuroticism, a risk factor for anxiety [Craske et al., 2009b], and as a predictor of the subsequent onset of anxiety disorders [Craske et al., 2011]. Individuals with, or at risk for, anxiety disorders appear to have difficulty inhibiting their fear to specific cues that should represent safety when they occur within a threatening context. By example, the anxious child on the playground who is bullied is not only fearful of the bully, but of the other children as well.

Direct evidence of deficits in inhibitory control is dependent on laboratory paradigms capable of teasing apart inhibition from excitation during fear learning, such as ‘conditional discrimination’
Fig. 2. Strategies for optimizing learning during exposure therapy.

Optimizing Exposure Strategies

We have targeted various methods of optimizing exposure therapy, derived from principles of fear extinction as well as inhibitory regulation. These are depicted in figure 2. Some but not all of these strategies will be described in the sections that follow; further details are provided by Craske et al. [2008a; 2012; 2014].

Violation of Expectancies

One approach is to design exposures in such a way that the experience maximally violates the negative, excitatory expectancies regarding the rate or frequency with which aversive outcomes occur [Rescorla and Wagner, 1972; Gallistel and Gibbon, 2000], or regarding the intensity of the aversive outcome [Davey, 1992], which in turn should enhance the development of inhibitory, non-threat expectancies. The greater the mismatch between what is expected to occur and what actually occurs, the greater the opportunity for learning [Rescorla and Wagner, 1972]. Thus, the goal is to provide experiences in exposure therapy that maximally disconfirms the individual’s threat expectancy. One method is to conduct exposure for periods of time that exceed the point at which the aversive event is expected to occur, which we implemented in a sample of acrophobic individuals [Baker et al., 2010]. Prior to beginning an exposure trial, participants indicated at what length of time in the height situation they fully expected an aversive outcome (i.e. falling). Over 2 days of exposure, participants were randomized to exposure trials at durations that specifically exceeded the durations at which they believed a negative outcome was likely to occur (disconfirmation) or exposure trials that discontinued before the duration when they believed the negative event was likely to occur (non-disconfirmation).

Thus, the evidence suggests that extinction is mediated by inhibitory learning/regulation and that individuals with anxiety disorders possess deficits in such inhibitory learning/regulation. Such deficits may serve as a risk factor for anxiety disorders, explaining the longevity of fears in disordered groups versus the transient nature of fears in non-disordered samples. At the same time, such deficits may render individuals with anxiety disorders in need of inhibitory learning and regulation during exposure therapy.

Given the behavioral data, one would expect anxious individuals to show deficits in vmPFC during extinction. Indeed, there is evidence for decreased orbitofrontal and medial PFC (including vmPFC / subgenual anterior cingulate cortex (ACC)) during extinction and at extinction retest in adults with PTSD and high trait anxiety relative to increased activity in those areas in healthy controls [e.g., Bremner et al., 2005; Milad et al., 2007; 2009; Rougemont-Bucking et al., 2011; Indovina et al., 2011].

Thus, the evidence suggests that extinction is mediated by inhibitory learning/regulation and that individuals with anxiety disorders possess deficits in such inhibitory learning/regulation. Such deficits may serve as a risk factor for anxiety disorders, explaining the longevity of fears in disordered groups versus the transient nature of fears in non-disordered samples. At the same time, such deficits may render individuals with anxiety disorders in need of means for augmenting inhibitory learning and regulation during exposure therapy.

We routinely design exposure trials to disconfirm expectancies regarding the temporal latency, frequency, and intensity of negative outcomes. In this approach, exposure tasks are designed to enhance new learning and are neither guided by premises of fear reduction nor to stay in the situation until fear declines. We establish what it is that the individual is most worried about (i.e. threat expectancy), and then design the exposure to disconfirm their expectancy, even if that means sustained fear activation (such as the person for whom repeated panic attacks may be optimal for disconfirmation of the expectancy of ‘going crazy’). Sometimes the expected outcome is an ‘inability to tolerate discomfort’ or an ‘inability to function’, in which case the exposure is designed to test out the ability to accomplish the exposure task despite high levels of fear, perhaps with the addition of tasks to test functioning (e.g., ask a question at a meeting when anxiety is elevated and record the answer in writing as a ‘test’ of functioning). In this model, focus upon fear reduction may ‘miss the target’ in terms of the expectancy to be disconfirmed. This approach ties exposure parameters directly to consciously stated expectancies for aversive events. As such, it overlaps with cognitive models of exposure as a vehicle for disconfirming misappraisals [Salkovskis et al., 2007].

Additional approaches described in the following do not rely upon conscious appraisals of expectancies.

In ‘deepened extinction’ [Rescorla, 2006], multiple fear conditional stimuli are first extinguished separately before being combined during extinction; joint presentation of 2 ‘predictors’ should elevate the expectancy of the US and thus allow another opportunity for learning. We found that such procedures reduced spontaneous...
ous recovery and reinstatement effects in a human conditioning/extinction study of healthy controls [Culver et al., 2014]. Notably, clinical applications have existed for some time in the form of implementing interoceptive exposure to feared bodily sensations (e.g., increased heart rate), first alone and then in combination with in vivo exposure to agoraphobic situations that also have previously undergone exposure alone [Barlow and Craske, 1994]. Another example would be exposure to one specific type of spider, then a second distinctly different spider, followed by exposure to both spiders at the same time. A third example would be exposure to an obsession (such as an obsession of stabbing a loved one), exposure to a cue that triggers obsessions (such as a knife in the presence of a loved one), followed by exposure to both the obsession and the knife in the presence of the loved one. As with the procedures for explicitly violating expectancies above, fear reduction is not a goal of this method of conducting exposures. Rather, the goal is to learn (explicitly or implicitly) that the aversive event occurs less often or is less intense than expected, achieved in this case through experience with more than one ‘predictor’ of such events.

Occasional reinforced extinction involves occasional CS-US pairings during extinction training (i.e. on 2 out of every 8 trials of extinction, the CS is repaired with the US), and has been shown to reduce rapid reacquisition effects [Bouton et al., 2004]. The benefits may derive from an expectancy violation effect in which the participant is less likely to expect the next CS presentation to predict the US because CS-US pairings have been associated with both further CS-US pairings and CS-no US pairings [Bouton et al., 2004]. Although, alternatives exist, in that the procedure of occasional reinforcement during extinction may enhance salience of the CS which in turn contributes to new learning about the CS [Pearce and Hall, 1980]. In a human conditioning study [Culver et al., 2014] we found that even though occasional reinforced extinction sustained fear arousal during extinction (i.e. no habituation), it attenuated the subsequent reacquisition of fear. The phenomenon of rapid reacquisition applies to situations of repeated aversive events, such as occurs in dangerous environments that lead to re-traumatization, but the approach of occasionally reinforced extinction is ethically prohibitive in such cases. Rapid reacquisition also applies to occasional social rejection in the case of social anxiety disorder and occasional panic attacks in the case of panic disorder. The clinical translation of occasional reinforced extinction would be the addition of occasional social rejections and ‘shame attacks’ in exposures to social situations, or the deliberate induction of panic attacks (e.g., by substances like yohimbine) in exposures to feared situations for panic disorder. As with deepened extinction, occasional reinforced extinction is not reliant upon violation of explicit expectancies. As such, deepened extinction differs from behavioral testing for the sake of disconfirmation of beliefs.

Cognitive restructuring strategies can impact threat expectancies. For example, ‘decatastrophizing’ may help to mitigate conditioned fear responding through reinterpretation of the intensity of the US [Lovibond, 2004]: in social anxiety disorder, the experience of rejection can be reinterpreted as less aversive (‘It wasn’t that bad’) or less indicative of additional rejection by others (‘Just because it didn’t work out with one person, doesn’t mean I can’t make other friends’). Cognitive strategies may also be helpful in generalizing extinction learning to additional stimuli, which can be beneficial because extinction is more stimulus-specific and does not generalize as readily to related stimuli as initial threat conditioning [Verhulst et al., 2006]. Cognitive interventions may facilitate this process by generating rules that link stimuli or by directing attentional resources to shared perceptual or categorical features (and making explicit what was learned during exposure). For example, following exposure to a ‘contaminated’ toilet in obsessive-compulsive disorder, Socratic questioning may assist the individual in noticing the similarities between the target stimulus (e.g., toilet) and other feared items (door handle, bathroom sink, etc.).

However, cognitive interventions may also have deleterious consequences if used in conjunction with some extinction strategies. Inasmuch as extinction relies on a violation of expectancy, using traditional cognitive restructuring of threat overestimation prior to exposure may reduce US expectancy and mitigate extinction learning. Similarly, reducing the aversiveness of the US through decatastrophizing prior to exposure may reduce any net change in associative strength, thereby mitigating extinction. Currently, we are preserving cognitive restructuring exercises for the time after exposure therapy, as a method for enhancing consolidation of learning, rather than beforehand when they could mitigate optimal mismatches between threat expectancies and outcomes.

**Stimulus Variability**

Varying the to-be-learned task enhances retention of learned non-emotional material [Schmidt and Bjork, 1992]. For example, random practice enhances long-term retention compared to blocked practice in which all trials of a particular task are completed prior to moving on to the next task [Shea and Morgan, 1979]. Variability is believed to enhance the storage capacity of newly learned information [Bjork and Bjork, 1992; 2006], to pair the information to be learned with more retrieval cues [Estes, 1955], and to generate a rule that captures the invariance among tasks [Schmidt and Bjork, 1992], which renders the information more retrievable at a later point in time. Variability can be explained by context retrieval models as well [Bouton, 1993], since variability is more likely to characterize encounters with phobic stimuli once exposure therapy is complete. Furthermore, by varying the stimulus, threat expectancies are likely to be elevated in some trials, thus providing more opportunities for disconfirmation and learning [Rescorla and Wagner, 1972], and attention to the CS is likely to be enhanced, which may facilitate learning [Pearce and Hall, 1980]. Indeed, we recently showed that attentional bias modification training designed to increase attention to the CS during extinction training lessened spontaneous recovery 1 week later in terms of associative measures of expectancy for the US and arousal measures of skin conductance (publication by Liao and Craske in preparation).

We found that variable stimuli resulted in less fear at follow-up behavioral avoidance testing in spider-fearful [Rowe and Craske, 1998a] and height-fearful samples [Lang and Craske, 2000], although a third study of contaminant anxiety showed trends only
[Kircanski et al., 2012]. Traditional exposure proceeds steadily from one hierarchy item to the next, with each item repeated a number of times until anxiety decreases. Instead, in variable exposure, exposure is conducted to items from the hierarchy in random order, without regard to fear levels or fear reduction, although usually beginning with the least anxiety-producing item to avoid treatment refusal. Thus, we routinely conduct exposure with varying stimuli, for varying durations, at varying levels of intensity, rather than continuing exposure in one situation until fear declines before moving to the next situation. Notably, such variability typically elicits higher levels of physiological arousal and subjective anxiety that does not apply to habituation during exposure [e.g., Lang and Craske, 2000], despite beneficial effects in the long term.

**Enhancing Retrieval of Inhibitory Learning**

**Consolidation Scheduling:** Basic science of learning argues that progressively increasing periods of time between learning trials enhance the retrievability of newly stored information [Bjork and Bjork, 1992]. In accord, we have shown that progressively increasing amount of time between exposure practices (such as 1 day, 4 days, 10 days) is more effective at follow-up than massed exposure in 2 studies of spider-fearful samples [Rowe and Craske, 1998a; Tsao and Craske, 2000]. This result speaks to the value of booster sessions following completion of exposure therapy, which we have shown to be a correlate of superior outcomes at 18 month follow-up [Craske et al., 2006].

**Retrieval Cues:** One option for enhancing retrieval of extinction learning and offsetting context renewal is to include retrieval cues (of the CS-no US association) during extinction training to be used in other contexts, once extinction is over [Brooks and Bouton, 1994; Vansteenwegen et al., 2006; Dibbets and Maes, 2011]. One risk of retrieval cues, however, is that they may become a safety signal (i.e., an indicator of the non-occurrence of the US) [Dibbets et al., 2008]. Proper use of retrieval cues evokes memory of the CS-no US non-threat expectancy without indicating absence of the US per se. In one study [Culver et al., 2011], we obtained weak effects of a retrieval cue (distinctive pen and clip board) upon context renewal. In another study, instructions to mentally restate what was learned during exposure (an instructional retrieval cue) had more robust effects on reducing context renewal [Mystkowski et al., 2006].

In the treatment of anxiety disorders, this approach prescribes that individuals carry cues (e.g., wrist bands) to remind them of what they learned during exposure therapy (as long as the cues do not become safety signals), or are prompted to remind themselves of what they learned in exposure therapy each time they encounter previously feared sensations or situations. However, these strategies are best employed as a relapse prevention skill. Using retrieval cues early in therapy, while the focus is on acquisition of extinction learning, may negatively impact progress as these cues can reduce the expectancy of the aversive event (and therefore mitigate expectancy violation effects).

**Scopolamine:** Given that context specificity of extinction is largely dependent on the hippocampus, strategies designed to down-regulate the hippocampus during extinction offer the potential to release context specificity, and thereby increase generalization of extinction learning. Indeed, animal studies have demonstrated that lesions of the hippocampus and pharmacological agents such as scopolamine that 'deactivate' impact hippocampal (and other brain regions) functioning reduce the context renewal of fear in rodents [Zelikowsky et al., 2012]. We are currently testing the effects of scopolamine upon long-term outcomes from exposure therapy for public speaking anxiety, in the same context as exposure and a novel context, manipulated using virtual reality.

Notably, the converse of generalization augmentation is desirable immediately following aversive events, when most paramount is context specificity in order to lessen the generalization of fear. The overgeneralization of fear that occurs following aversive events, especially in at-risk individuals, may be partially mediated by down-regulation of the hippocampus via excessive cortisol levels. Notably, glucose spares hippocampal functioning. Following from basic animal research [Minor and Saade, 1997], we found that administration of glucose immediately following conditioning enhanced contextual conditioning relative to placebo [Glenn et al., 2014]. We are currently exploring the potency of glucose ingestion immediately following naturally occurring stressors.

**Positive Valence Training:** CS valence, or the degree of fondness/aversion towards the CS, may contribute to return of defensiveness responding after reinstatement [e.g., Hermans et al., 2002]. In addition to acquiring an expectancy of the US, the CS develops a negative valence during threat conditioning [e.g., Hermans et al., 2005; 2002]. Acquired valence of the CS is a type of evaluative learning [De Houwer et al., 2005] and is more resistant to extinction than is expectancy learning [Dirikx et al., 2004; Hermans et al., 2002]. Hermans and his colleagues [Hermans et al., 2005] found a positive association between CS negative valence at the end of extinction and degree of subjective fear following unpaired US presentations (i.e., reinstatement). We replicated these effects (Zbozinek et al., under review). These findings raise the possibility that increasing the positive valence towards the CS by the end of extinction (or exposure therapy) may reduce the effects of reinstatement. We have demonstrated such effects in 2 separate studies. The first (Zbozinek et al., under review) was a laboratory conditioning study, in which we randomized participants to positive imagery training, a paradigm developed by Emily Holmes and colleagues to induce positive affect, or positive verbal training, following conditioning but before extinction. Positive imagery training led to more positive affect, more positive valence towards the CS by the end of extinction (or exposure therapy) may reduce the effects of reinstatement. Furthermore, regardless of group assignment, the more positively the CS was evaluated at the end of extinction, the less fear reinstatement, thus replicating previous findings. This study relied upon a generic mood induction technique to raise positive affect, which then spilled over to result in more positive evaluations of the CS. In clinical application, this would highlight the potential of general positive mood induction techniques prior to exposure therapy. In a second study (Dour et al., under review), we evaluated positive information about the phobic stimulus, presented at the end of the first day of 2 days of in-vivo exposure to spiders. The
positive information was presented via an animation in which spiders were portrayed as essential to the ecosystem, hardworking, wise, and kind. Spider-fearful individuals assigned to the positive information video about spiders reported less fear of spiders after exposure therapy and reinstatement (via a negative video of a spider consuming a mouse), and more approach towards a spider in a behavioral avoidance test. This study suggests that information designed to improve positivity towards the phobic stimulus may be of benefit.

**Therapeutic Strategy for Enhancing Inhibitory Regulation**

Social neuroscience has identified another strategy for enhancing inhibitory regulation which involves linguistic processing or affect labeling. Affect labeling may work to augment associative inhibitory processes within extinction or may work in an independent but complementary manner to extinction learning. A number of studies have shown that linguistic processing activates a region of the cortex, the right ventrolateral prefrontal cortex that reduces activity in the amygdala, thereby attenuating anxious responding [Lieberman et al., 2007]. It appears that engaging the executive functioning cortical areas of the brain works to dampen the limbic system activity. Furthermore, we have shown that this pathway is disrupted in anxious samples and particularly socially anxious individuals, relative to healthy controls [Burklund et al., 2015]. While the ventrolateral PFC was activated during affect labeling, the amygdala was not down-regulated in socially anxious individuals, and was even elevated for socially anxious individuals who were also depressed. This finding suggests that anxious and depressed individuals show deficits in an implicit form of emotion regulation that parallels the deficits in extinction learning and inhibitory pathways underlying extinction learning described earlier.

As such, anxious and depressed individuals may benefit from training in affect labeling to the extent that such training can improve inhibitory regulation of emotion. In 3 studies, we have shown benefits of affect labeling as individuals are exposed to feared stimuli. Tabibnia et al., [2008] found that repeated evocative spider images paired with word labels, negatively valenced and irrelevant to the images (e.g., bomb and war), produced a greater reduction in subsequent skin conductance response (SCR), 1 week later, than unpaired images. Kircanski et al. [2012] found added benefits of affect labeling in a sample of individuals with spider phobias as they underwent exposure therapy. In comparison to cognitive reappraisal of thoughts, distraction, and exposure alone, affect labeling during exposure was found to reduce skin conductance and to increase approach behavior at 1 week follow-up in a context different than the exposure context [Kircanski et al., 2012]. In addition, the more affect labeling words that represented fear, anxiety or other negative emotions, the better the outcomes at follow-up. Furthermore, we have recently shown benefits from affect labeling during public speaking for socially anxious individuals [Niles et al., 2015]. Participants assigned to the affect labeling group, especially those who used more labels during exposure, showed greater reduction in physiological activation at follow-up behavioral testing than participants in the control group. No effect was found for self-report measures. In addition, greater emotion regulation deficits at baseline predicted more benefit from exposure combined with affect labeling than exposure alone on physiological arousal. Together, these data suggest that affect labeling, as opposed to more traditional cognitive therapy which attempts to change the content of appraisals, can improve physiological and behavioral outcomes from exposure. We routinely ask clients to state their emotional responses, without attempting to change their emotional responses, in the midst of exposure. We are currently testing out the use of affect labeling during brief exposure to reminders of a trauma, through script driven imagery procedures, for posttraumatic stress disorder and the associated neural effects.

**Summary of Inhibitory Model of Exposure Therapy**

Our approach to exposure therapy follows the premises of inhibitory learning and inhibitory regulation. First, the therapist and client decide on the specific goal of the practice, followed by clarification of the threat expectancy, or the feared outcome of engaging in the task. Exposures are then designed in such a way and proceed until a given anticipation or expectation is violated, which may begin with conditions that do not optimally violate those expectancies and then progress to conditions which do. In addition, exposu-
re includes ‘inhibitory learning enhancement and inhibitory regulation enhancement strategies’ with deepened extinction (or exposure to multiple feared cues), occasionally reinforced extinction (or occasional exposure to aversive outcomes), weaning from safety signals, stimulus variability, consolidation scheduling, retrieval cues, multiple contexts, positive valence training, and affect labeling. Scopolamine may provide an additional augmentation strategy. The exposure plan showed in figure 3 includes some of the augmentation strategies for a socially anxious individual who fears and avoids parties; the exposure task is to attend a party, and given a threat expectancy of being rejected for expressing an opinion about a difficult topic. The first exposure plan is to discuss politics at a social event. The practice includes throwing out a safety behavior of avoiding eye contact, and deepened extinction by adding in strategies designed to increase body temperature (wearing a warm jacket) since the client expects that showing signs of anxiety will also lead to rejection by others. The client also plans to label his emotions just beforehand. Immediately before the practice, the client rates threat expectancy.

Following completion of exposure therapy, therapists aid clients in discussing the violation of expectancies to consolidate the new learning, as shown in figure 4.

**Conclusion**

Framing exposure within a modern learning theory perspective holds numerous advantages, including providing a parsimonious explanation for shared elements of traditional exposure (or behavioral experiments), while simultaneously explaining their shortcomings. In addition, it ties clinical research to the wealth of research on learning theory in animal and human populations. Further, it holds promise for improving the efficacy of exposure-based procedures through selective targeting of associative learning mechanisms. Associative learning theories provide a parsimonious explanatory model from which to situate exposure processes. However, additional translational research is needed to further elucidate the optimal conditions necessary for enhancing inhibitory regulation and the precise methods for implementing these strategies in routine clinical care.

**Disclosure Statement**

The author declares that there is no conflict of interest concerning this paper.

---

**References**


**Optimizing Exposure Therapy for Anxiety Disorders**
Estes WK: Statistical theory of distributional phenomena

Garcia-Palacios A, Botella C, Hoffman H, Fabregat S:


Potter J, Dippa-Perea L, Melendez LM, Quirk GJ. Inhibition of fear extinction with hippocampal-infralimbic BDNF. Science 2010; 1288–1290.


Quirk GJ. Memory for extinction of conditioned fear is long-lasting and persists following spontaneous recovery. Learn Mem 2002; 9: 402–407.


Ricker ST, Bouton ME. Reacquisition following extinction in appetitive conditioning. Anim Learn Behav 1996; 24: 423–436.


