Treatment of Avoidance Behavior as an Adjunct to Exposure Therapy: Insights from Modern Learning Theory

Michael Treanor[1*]

And

Tom J Barry[2,3]

[1]University of California, Los Angeles

[2]Institute of Psychiatry, Kings College London
[3]Experimental Psychopathology Lab, Department of Psychology, The University of Hong Kong

[*]Correspondence should be addressed to

Michael Treanor, Ph.D.
Department of Psychology
University of California, Los Angeles
1285 Franz Hall
Box 951563
Los Angeles, CA 90095
mtreanor@psych.ucla.edu

Abstract

Pathological avoidance of benign stimuli is a hallmark of anxiety and related disorders, and exposure-based treatments have often encouraged the removal of avoidance, or safety behaviors, due to their negative effects on extinction learning. Unfortunately, empirical evidence suggests that avoidance behaviors can persist following treatment, and the mere availability of avoidance behavior can be sufficient to renew fear following successful extinction learning. The present paper critically examines the function of avoidance behavior through the lens of modern learning theory, and speculates on novel behavioral and pharmacological strategies for targeting avoidance as an adjunct to current evidence-based treatments.

Avoidance has long held a central role in theories regarding the genesis and maintenance of anxiety disorders. For example, Mowrer (1951) conceptualized avoidance as maintained through negative reinforcement resulting from anxiety reduction. More recently, avoidance has been conceptualized as being driven by expectation that a stimulus will lead to an aversive outcome (Declercq, De Houwer, & Baeyens, 2008). In both instances, avoidance becomes pathological when performed in response to relatively benign stimuli.

Although avoidance has been important in theories of anxiety, translational research has largely focused on other Pavlovian processes, such as extinction learning, as the principal mechanism of treatment for anxiety disorders (e.g., exposure therapy). The implicit assumption has been that avoidant behavior would decrease as individuals learned that a threatening stimulus (conditional stimulus or CS) no longer predicted an aversive outcome (unconditional stimulus or US). That is, following extinction training, and the repeated presentation of the CS in the absence of the US, there would no longer be any need to avoid the CS. However, empirical evidence suggests that avoidance behavior can persist following extinction (Rodriguez-Romaguera, Greenberg, Rasmussen, & Quirk, 2016; Solomon, Kamin, & Wynne, 1953), and the availability of avoidant behavior can renew fear even following successful extinction learning. For example, Vervliet and Indekeu (2015) conditioned avoidance behavior (a button press prevented a shock during a CS presentation) and then conducted extinction training where the avoidance behavior was not available. Self-reported fear and physiological arousal to the CS decreased during the extinction phase, however, simply making the avoidant response available at a later test phase when the CS was presented again caused

fear to return to the CS. Similar results have been obtained in rodents (Bravo-Rivera, Roman-Ortiz, Montesinos-Cartagena, & Quirk, 2015).

This presents obstacles to evidence-based interventions based on extinction such as exposure therapy. In exposure-based treatment clients are often encouraged to refrain from avoidant behavior (e.g., use of anxiolytic medication, compulsive behaviors, having a “safe” person). However, the above evidence suggests that avoidance behavior may persist, and the mere availability of an avoidant response may be sufficient to renew fear following treatment. This may represent one reason patients relapse following exposure therapy (Ginsburg et al., 2014).

The reason for the deleterious impact of avoidant behavior availability following successful extinction or exposure remains unclear. One possibility is that removing avoidant behavior during extinction represents a context shift such that it differs from both the original context in which fear was acquired as well as other contexts that might be encountered after extinction/exposure. As such, when the avoidance response is available again, this represents another context shift from extinction, and fear is then renewed in the same way that it might if extinction had taken place in a different physical context/environment (Vansteenwegen et al., 2005; Vervliet & Indekeu, 2015). Regardless, these findings suggest that it may be important to examine the treatment of avoidance behavior as an adjunct to exposure-based procedures in order to mitigate renewal of symptoms (Vervliet & Indekeu, 2015).

The present paper critically examines the treatment of avoidant behavior through the lens of modern learning theory. Through examination of the various functions avoidant behavior may serve in associative learning processes, as well as its

neurobiological substrates, we aim to highlight novel behavioral and pharmacological interventions that may serve as useful adjuncts to traditional evidence-based strategies for anxiety and related disorders. In addition, given the dearth of evidence elucidating the mechanisms responsible for the return of fear following treatment as a result of the availability of avoidance behavior, we conclude with concrete recommendations for future research.

Avoidant Behavior as a Conditional Inhibitor

Extinction of conditional inhibition

Extinction learning is one of the presumed mechanisms of exposure therapy (Craske et al., 2008; Scheveneels, Boddez, Vervliet, & Hermans, 2016) and operates via error correction mechanisms, such that the associative strength of a CS is updated when the US does not occur. During learning, the greater discrepancy between what is predicted and what actually occurs, the larger the amount of associative change that takes place (Rescorla & Wagner, 1972). Conditional stimuli that predict the occurrence of a US are known as “conditional excitors” whereas stimuli that directly predict the absence of the US are “conditional inhibitors”. During extinction training, in which a conditional excitor is repeatedly presented in the absence of the US, the concurrent presence of conditional inhibitors decreases the expectation that a US will occur, resulting in less expectancy violation, and therefore negatively impacts extinction learning (Lovibond, Chen, Mitchell, & Weidemann, 2013; Lovibond, Mitchell, Manard, Brady, & Menzies, 2009; Rescorla, 1969).

Avoidant behaviors, or “safety behaviors”, have often been discussed in terms of conditional inhibition (Krypotos, Effting, Kindt, & Beckers, 2015). For example, the use

of benzodiazepines in panic disorder, washing one's hands in obsessive-compulsive disorder, and a combat veteran sitting with his back to a wall in a restaurant are all examples of avoidant behaviors that may function as conditional inhibitors as they are directly associated with the decreased likelihood of the US occurring (See Figure 1 for a graphical representation of the relationship between a CS+, conditional inhibitor, and a US). Importantly, despite functioning as a conditional inhibitor, the availability of avoidance behavior may still become a contextual feature and lead to context renewal (Vervliet & Indekeu, 2015). That is, when the avoidance behavior is available following treatment, this may represent a context shift from exposure procedures when avoidance was prohibited, and results in a return of fear. The implication would be that allowing some avoidance behavior during exposure may be beneficial to reduce subsequent context renewal, although the deleterious impact of avoidance behavior (e.g., conditional inhibitors) on extinction learning represents a significant problem. Thus, conditional inhibition has to be reduced, or the negative effects of conditional inhibitors on extinction learning needs to mitigated, prior to allowing avoidance behaviors that function as inhibitors during exposure therapy. Below, we discuss specific treatment approaches for targeting conditional inhibitors as an adjunct to exposure therapy.

The traditional paradigm for developing conditional inhibition is to pair a neutral stimulus (B) with an excitatory stimulus (A) without reinforcement (e.g., A+ then AB-). The resulting decrease in associative strength gradually transforms the previously neutral stimulus into an inhibitor. For example, engaging in compulsive behavior (neutral stimulus) when one has obsessive thoughts (conditional stimulus) gradually transforms the compulsive behavior into a conditional inhibitor when the US doesn't occur.

However, dominant learning models suggest that presenting a conditional inhibitor by itself should result in a gradual loss of inhibition, and may offer one potential strategy for targeting avoidance behavior. Rescorla & Wagner (1972) conceptualized change in associative strength as a function of the total amount of learning a US can support (λ) minus the sum of the associative strength of all the stimuli present on a given trial (ΣV). Let us assume a negative associative strength of a conditional inhibitor of -.5. Presenting it alone, in the absence of another CS or US should result in a net positive amount of associative change ($\lambda - \Sigma V$ becomes $0 - [-.5]$) that will gradually eliminate inhibition (Zimmer-Hart & Rescorla, 1974). For example, an individual with obsessive-compulsive disorder may be asked to wash her hands compulsively in the absence of touching a contaminated surface while someone with panic disorder may be asked to take a benzodiazepine at times when he is not anxious. Although this is consistent with dominant learning models, numerous animal studies have failed to find any loss of inhibition after repeatedly presenting a conditional inhibitor in isolation (e.g., DeVito & Fowler, 1987).

However, in a study of human contingency learning, Melchers, Wolff and Lachnit (2006) argued that one *could* produce extinction of conditional inhibition depending on the nature of the US. The authors argue that traditional Pavlovian procedures use unconditional stimuli that only vary unidirectionally. For example, one is either shocked or not shocked in conditioning and extinction experiments. However, the Rescorla-Wagner model's assumption that inhibition is the opposite of excitation would necessitate that the US can take on values less than zero. When the US can only vary in one direction, a conditional inhibitor predicts the non-occurrence of the US and there is no

discrepancy, or extinction learning, when it is presented alone without the US. However, when the US can take on both positive and negative values, then presenting a conditional inhibitor in isolation can still lead to expectancy violation.

In the Melchers et al. (2006) study, participants were divided into two groups tasked with determining whether a fictional individual's hormone levels would rise (US) based upon consumption of certain foods (CS). In one group, hormone levels could only rise or remain the same (unidirectional group), whereas in the other group hormone levels could rise, remain the same, or decrease (bidirectional group). Using standard paradigms for developing inhibition the authors demonstrated that you could reduce conditional inhibition through non-reinforced presentations of the inhibitor but only in the group in which the US was allowed to vary bidirectionally (Melchers et al., 2006). Subsequent studies have replicated and extended this result (Baetu & Baker, 2010; Lotz & Lachnit, 2009).

If extinction of conditional inhibition is possible with a bidirectional US, then the question becomes whether clinical disorders meet this criterion. In social anxiety disorder, exposures are often tailored towards interpersonal interactions. The feared consequence or US may be rejection, but the US could be conceptualized as lying on a continuum ranging from approval, to a neutral response, to rejection. Thus, social anxiety disorder may be one instance in which a perceived US can vary bidirectionally, and therefore avoidant behaviors that function as conditional inhibitors may be amenable to unreinforced exposure. However, other disorders, such as panic disorder, may entail unconditional stimuli that are best characterized as unidirectional. A person may only suffer a heart attack or not as a result of a rapid heartbeat. A task for future research is to

determine whether or not unconditional stimuli in anxiety disorders are best characterized as unidirectional or bidirectional, and the degree to which conditional inhibitors in these disorders are amenable to extinction learning via presentation of the inhibitor in the absence of the excitatory CSs or US. In addition, it may be important to examine the bi-directional nature of unconditional stimuli across disorders (e.g., eating disorders) to further elucidate potential translational applications. Although extinction of inhibition represents an exciting possibility, considerable additional research is needed to further explore its translational applicability.

Aversive learning and counter-conditioning

Vervliet and Indekeu (2015) suggest that the availability of avoidance behavior following extinction may result in fear renewal partly because it is a context shift from extinction (where avoidant behavior was not available). They argued that presenting the avoidant behavior occasionally during extinction, and therefore making it a feature of the extinction context, might facilitate generalization of extinction learning by making it more similar to other contexts one encounters outside of extinction/exposure. However, presenting a conditional inhibitor during extinction, even sparingly, might still negatively impact extinction learning on a given trial.

Conditional inhibitors mitigate extinction learning because of their negative associative “charge” that reduces expectancy violation. It follows that if this negative charge can be reversed, such that the inhibitor is now itself associated with an aversive event, then this may reduce its detrimental effects on extinction. Indeed, initial results in rodents demonstrated that pairing the inhibitor with a reinforcer reduced conditional

inhibition (Holland, 1984; Zimmer-Hart & Rescorla, 1974), although some inhibitory strength may remain (Pearce &Wilson, 1991).

Unfortunately, this may have limited clinical utility. It is certainly not ethical to present the US (e.g., a trauma) alongside a conditional inhibitor in most anxiety disorders. However, in social anxiety disorder the US (e.g., rejection) can occur without detrimental results to the patient. For example in occasional reinforced extinction the CS is intermittently paired with the US to enhance extinction learning (Bouton, 2004). Clinically, this most often takes the form of "shame attacks" in social anxiety disorder where the individual engages in behavior that has a high likelihood of rejection (e.g., asking strangers if the earth revolves around the sun) during an exposure exercise. Allowing the individual to engage in avoidant behavior, such as hiding signs of anxiety, during occasional reinforced extinction may represent one strategy for altering the inhibitory strength of avoidance behaviors that function as conditional inhibitors. Although this strategy may be useful for social anxiety disorder, it still requires the presentation of an aversive US, and it is unlikely to be useful for other anxiety disorders.

Timing issues and attentional redirection

If allowing avoidance behavior during extinction is necessary to reduce return of fear than it is important to consider the optimal timing of avoidance behavior during extinction, as well as strategies to mitigate the deleterious effects of conditional inhibitors on expectancy violation. Rescorla & Wagner (1972) suggest that the largest changes in associative strength will occur in the early phases of extinction learning, as the CS+ is still a strong predictor of the US. During initial training, there is a large discrepancy between the predictive strength of the CS and the non-occurrence of the US. However, as

extinction proceeds this discrepancy is reduced as the CS-noUS relationship becomes stronger.

Inasmuch as there are smaller changes in associative strength in the later phases of extinction, then the presence of a conditional inhibitor will have less of a negative effect on extinction learning. This might represent the optimal time to combine the availability of avoidance behavior with extinction learning. The occasional availability of avoidant behavior may allow it to become a feature of the extinction context, thereby serving as a retrieval cue for extinction if it is available at a later time, while the infrequent presence of conditional inhibitors, combined with their use only during the later phases of extinction, will mitigate any negative effects on expectancy violation. Clinically, this may entail the availability of avoidance behaviors that function as conditional inhibitors in the last several sessions of exposure therapy, and then only during a select few trials, as a relapse prevention measure.

In addition to error-correction, associative change is enhanced by CS salience. Dominant learning models suggest that stimuli "compete" for changes in associative learning, such that the more salient stimuli acquire the largest amount (Mackintosh, 1975; Pearce & Hall, 1980, Rescorla & Wagner, 1972). Thus, in trials with compound stimuli, manipulations that enhance attention to one stimulus (e.g., A) at the expense of the other (e.g., B) may allow the target stimulus (A) to acquire the greater amount of associative change. When presenting a conditional inhibitor alongside a conditional excitor, enhancing the attentional salience of the excitor may allow it to accrue the bulk of associative change.

In addition, the impact of a conditional inhibitor on extinction learning may be further reduced through attentional manipulations. Learning models often employ a summed error term (ΣV) that determines the overall amount of associative change that occurs on a given trial. For example, if a conditional excitor and inhibitor have the same amount of associative strength then they will cancel each other out and result in no learning during extinction. However, the more salient a stimulus is, the more its current associative strength is likely to contribute to the summed error term. Reducing the salience of a conditional inhibitor may reduce the degree to which its inhibitory strength contributes to this summed error term, and therefore could mitigate the negative effects of conditional inhibitors during extinction. Clinically, this may take the form of explicit instructions to focus on the CS+ at the expense of the inhibitor. For example, in exposure therapy for panic disorder the patient may be provided with explicit instructions to attend to the sensations of a rapid heartbeat (CS+) while ignoring the anxiolytic medication in her purse (conditional inhibitor). However, additional research is necessary to explore these possibilities, along with examination of feasible treatment strategies for manipulating attention (e.g., attentional bias modification, explicit instructions).

Avoidance Behavior as Occasion Setting

Although avoidance behavior may often function as a conditional inhibitor, it may modify CS-US relationships in other manners as well. Modulatory stimuli, or occasion setters, are stimuli that are not directly associated with the US, but "set the occasion" for whether or not the CS will lead to the US (Holland, 1989). For example, in panic disorder, an individual may fear being alone with a rapid heartbeat. Although being alone is not directly related to the US (e.g., being alone is not predictive of a heart attack), the

individual may fear that a rapid heartbeat is more likely to lead to a heart attack if they are alone. Avoidance behavior has often been discussed in terms of negative occasion setting (Declerq & De Hower, 2008; De Hower, Crombez, & Baeyens, 2005; Kryptotos et al., 2015; See Figure 2 for a graphical representation of the relationship between a CS+, occasion setter, and a US). Occasion setting is more likely when the behavior or stimulus precedes the CS or is a less salient feature (e.g., the context; Holland, 1986). In anxiety disorders, typical avoidance behaviors that may function as negative occasion setters include cellular phones or "safe" individuals in panic disorder and preparing conversation topics ahead of time in social anxiety disorder.

As with conditional inhibition, one strategy may be to present the occasion setter alone, thereby extinguishing its modulatory properties or any additional association it might have with the US. Unfortunately, several experimental investigations in animals and humans have demonstrated that this strategy fails to result in extinction of occasion setting (Holland, 1989; Rescorla, 1986; van Vooren, Franssen, Beckers, Hermans, & Baeyens, 2012). However, reversing the contingencies of occasion setting, such that a negative occasion setter that predicted a CS would not lead to a US now predicts the CS will lead to the US, does successfully extinguish occasion setting (Rescorla, 1986). Unfortunately, similar to counter-conditioning conditional inhibition, reversing the contingencies of a negative occasion setter may have limited clinical applicability beyond social anxiety disorder.

Reducing the ability of avoidant behavior to renew fear may require occasionally presenting the avoidant behavior during extinction, thereby making it a feature of the extinction context, although this may simultaneously reduce expectancy violation and

mitigate extinction learning. We previously suggested several behavioral strategies for reducing these deleterious effects when the avoidant behavior functions as a conditional inhibitor (e.g., attentional modulation, allowing the avoidance behavior only during the later phases of extinction, etc.). Indeed, similar strategies should be examined in reducing the impact of avoidance behaviors that function as negative occasion setters, as the extant evidence suggests that occasion setters can similarly impact error-correction and conditional responding (Morel & Holland, 1993).

Pharmacological interventions for Conditional Inhibitors and Occasion Setters

There has been increased interest in pharmacological agents that may impact associative learning processes as adjuncts to traditional behavioral interventions (Quirk & Mueller, 2008). However, we are unaware of any research examining pharmacological treatment of avoidance behavior in the context of extinction and exposure procedures. Thus, the following section examines potential molecular pathways that may serve as useful targets for pharmacological agents in the treatment of conditional inhibition and occasion setting.

As mentioned previously, presentation of a conditional inhibitor in isolation (e.g., extinction of inhibition) may represent one behavioral strategy for reducing the impact of avoidance behavior on renewal of fear when the US varies bi-directionally. Neurobiological models of traditional extinction implicate neuroplasticity and de novo protein synthesis in the consolidation of the extinction memory (Quirk & Mueller, 2008). For example, a wealth of research has demonstrated the importance of the N-Methyl-D-aspartate (NMDA) receptor in extinction learning (Myers & Davis, 2007). Indeed, NMDA receptor agonists, such as d-cycloserine (DCS) have been found to enhance

extinction learning in both animals and humans (Davis, Ressler, Rothbaum, & Richardson, 2006), and have shown promise as adjuncts to exposure-based treatments (Guastella et al., 2008).

Importantly, NMDA receptor activity is not confined to fear extinction, and is an important component of a variety of memories including conditional inhibition (Foilb, Flyer-Adams, Maier, & Christianson, 2016). If NMDA receptor activity underlies both conditional inhibition and extinction learning then it is possible that DCS may be useful during extinction of inhibition. Extinction of inhibition requires presentation of the inhibitor in isolation. For example, an individual with post-traumatic stress disorder may repeatedly check his household locks while an individual with obsessive-compulsive disorder may wash her hands repeatedly. A key to extinction of inhibition is presenting the conditional inhibitor in the absence of both the CS+(stimulus that predicts the US) and US. Thus, in the examples above the patient would only engage in the avoidance behavior when they are not "triggered" or confronted with a CS+. If pharmacological agents, such as DCS, also enhance extinction of inhibition, then the amount of time dedicated to exposure to the conditional inhibitor in isolation can be reduced, while simultaneously strengthening the consolidation of extinction of inhibition. However, there may also be risk in attempting to enhance the consolidation of extinction of inhibition, as any failure in extinction may paradoxically result in strengthening conditional inhibition.

As discussed previously, reducing the impact of avoidance behavior on extinction generalization may occasionally require presenting the conditional inhibitor in compound with the CS+ during exposure therapy. However, presenting a conditional inhibitor also

reduces expectancy violation during a given trial and mitigates extinction learning. Pharmacological agents that enhance the excitatory strength of the CS+ may reduce the deleterious impact of avoidance behavior on extinction by magnifying expectancy violation despite the presence of a conditional inhibitor during an extinction trial.

The neurobiological substrates of error correction mechanisms have been linked to dopamine signaling, such that increased dopaminergic activity is associated with surprise or expectancy violation (e.g., Steinberg, Keiflin, Boivin, Witten, Deisseroth & Janak, 2013). Although originally associated with reward, there is increased evidence that dopamine signaling plays a role in aversive conditioning and extinction as well (Haaker et al., 2013; Hikind & Maroun, 2008; Holtzman-Assif, Laurent, & Westbrook, 2010; Lissek, Glaubitz, Wolf, & Tegenthoff, 2015; Mueller, Bravo-Rivera, & Quirk, 2010; Yuan Li & McNally, 2014).

During normal extinction, the non-occurrence of the US should produce prediction error for the CS+ but not a conditional inhibitor (Schultz, 2007). Indeed, dopamine neurons demonstrate reactivity during US omission to a CS+, but not during the presentation of a conditional inhibitor or conditional inhibitor combined with a conditional excitor (Tobler, Dickinson, & Schultz, 2003). The latter case represents a lack of prediction error when the CS+ is presented in compound with a conditional inhibitor. Thus, increasing dopaminergic activity during extinction with a CS+/conditional inhibitor compound may enhance the molecular substrates of expectancy violation, mitigating any negative effects of conditional inhibition on extinction learning. Unfortunately, we are unaware of any studies that have directly examined this possibility. However, amphetamine sensitization (which increases dopamine levels) reduces the effect of a

conditional inhibitor on a conditional excitor (Shiflett, Riccie, & DiMatteo, 2013 but see Harmer & Phillips, 1999), and suggests that dopamine agonists might offer protection against the negative effects of avoidance behavior during exposure therapy.

In addition to dopamine, opioid receptors play a role in error correction mechanisms. Opioid release modulates Pavlovian learning such that excitatory conditioning corresponds with reduced opioid release whereas extinction learning is facilitated by increased opioid activity (Myers & Davis, 2007; Yuan Li & McNally, 2014). However, as with dopamine agonists, a key question is how modulating opioid activity affects extinction learning in combination with a conditional inhibitor. As discussed previously, during presentation of an excitatory/inhibitory compound, the learning history of each stimulus affects responding to the compound (e.g., inhibition can reduce expectancy violation). Blocking is a related cue competition phenomenon where a well-established conditional excitor (A+) prevents conditioning to a new stimulus (B) when they are presented in compound (AB+). This occurs because the occurrence of the US is already well predicted by stimulus A, so no new learning accrues to stimulus B. However, modulating opioid activity reduces blocking, and allows conditioning to stimulus B (Iordanova, McNally, & Westbrook, 2006). This may be tantamount to a well-conditioned inhibitor (that strongly predicts the non-occurrence of the US) reducing extinction learning when combined with a conditional excitor, and suggests that targeting opioid receptors is a promising intervention. Indeed, recent evidence suggests that modulating opioid activity reduces the impact of a conditional inhibitor on a conditional excitor in rats (Laurent, Wong, & Balleine, 2015).

Unfortunately, pharmacological agents may have limited utility when targeting occasion setting. For example, inasmuch as presenting an occasion setter alone does not appear to result in extinction learning (Holland, 1989; Rescorla, 1986), pharmacological enhancement of memory consolidation (as discussed above) will not be helpful. In addition, hippocampal lesions and pharmacological disruption of hippocampal activity reduce the acquisition of occasion setting (Meyer, Putney, & Bucci, 2015; Yoon, Graham, & Kim, 2011), but it is unclear how this would be leveraged in clinical situations, as the avoidant behavior (i.e., negative occasion setting) would have already been acquired.

However, if, as Vervliet and Indekeu (2015) suggest, the availability of avoidance behavior following treatment can renew fear because it represents a context shift from extinction, then pharmacological strategies that may reduce the contextual gating of extinction learning may be helpful. The contextual nature of extinction learning has been linked to increased hippocampal activity during extinction (Hermann, Stark, Milad, & Merz, 2016; Holland & Bouton, 1999). In an elegant study, rats received systemic injections of the anticholinergic scopolamine. The hippocampus is rich in cholinergic receptors (Yi et al., 2015), and therefore "impairing" hippocampal activity may reduce the ability of the hippocampus to bind extinction learning to a particular contexts. Indeed, rats treated with scopolamine generalized extinction learning to a novel context, whereas vehicle treated rats demonstrated increased fear consistent with context renewal (Zelikowsky, et al., 2013). This may suggest that pharmacological agents that disrupt hippocampal activity during extinction may allow learning to generalize to new contexts, and may represent one strategy to mitigate any contextual specificity when avoidance

behavior is prevented during extinction or exposure procedures. Of course, one potential concern with combining extinction and pharmacological agents is that the drug may create an "internal context" that precipitates context renewal when the individual is tested at a subsequent time point in the absence of the drug (Bouton, 2004).

Summary

Despite the efficacy of extinction-based procedures (i.e., exposure therapy) in the treatment of anxiety disorders, pathological avoidance may still persist, and empirical evidence suggests that the mere availability of avoidance behavior can renew fear. The present paper attempted to elucidate novel behavioral and pharmacological interventions for avoidance behavior derived from learning theory and neurobiology. A few key findings emerged. First, consistent with previous discussions (Krypotos et al., 2015) avoidance behavior may serve various functions including conditional inhibition and negative occasion setting. Accurately determining whether a given avoidant behavior is functioning as a conditional inhibitor or negative occasion setter may be important, as each may require different behavioral and pharmacological strategies. Unfortunately, typical laboratory-based paradigms for assessing inhibition or occasion setting (e.g., tests of summation or retardation) will be difficult to carry out with actual avoidance behavior in clinical settings. Translational interventions may have to rely on detailed functional analyses in order to determine the function of an avoidant behavior. For example, if the behavior directly predicts the non-occurrence of the US, rather than simply modulating the likelihood that a CS will lead to a US, then it is most likely a conditional inhibitor.

Second, the type of intervention may differ depending on how avoidant behavior increases risk for renewal of fear. If removing avoidance behavior during extinction

represents a context shift (Vervliet & Indekeu, 2015), then allowing avoidance behavior during extinction may be necessary. However, this will require behavioral or pharmacological methods to reduce any negative effects of avoidance behavior on expectancy violation.

Each of these possibilities requires further empirical investigation. Below, we offer several concrete recommendations for future research.

1. Determine the degree to which a bidirectional US facilitates extinction of inhibition, and the extent to which anxiety disorders correspond to this criterion. Although several studies in human contingency learning have demonstrated this possibility, it is important to replicate these results in human fear conditioning with more externally valid types of unconditional stimuli. For example, standard laboratory conditioning and extinction paradigms can be combined with relevant unconditional stimuli that either vary bidirectionally or unidirectionally (e.g., breathing occlusion for panic patients or insults for individuals with social anxiety disorder; Lissek et al., 2008). Employing a conditional inhibitor in these paradigms will allow examination of extinction of inhibition depending upon the nature of the US.
2. Examine the impact of attentional manipulations (e.g., explicit or implicit training in attending to the CS+) on extinction alongside a conditional inhibitor and subsequent return of fear in both healthy controls and individuals with anxiety disorders.
3. Examine effects of pharmacological agents, such as dopamine agonists and opioid agonists/antagonists, on extinction of a CS+ in combination with a conditional

inhibitor. This may allow the avoidant behavior to become a feature of the extinction context, thereby reducing return of fear when it is made available during a subsequent test, while simultaneously reducing any negative effect on expectancy violation during extinction.

In examining the treatment of avoidance behavior alongside exposure procedures, it will be necessary to employ a more nuanced view of avoidance behavior derived from modern learning theory. For example, the extent to which a given avoidance behavior may negatively affect extinction learning may be dependent upon whether it is a conditional inhibitor or occasion setter, as well as the overall ratio of inhibition to excitation (i.e., stronger inhibitors will have a greater impact on extinction). Unfortunately, many studies have failed to take into account this degree of complexity, which may partially explain mixed results regarding whether or not avoidance is detrimental to exposure processes (Meulders, Van Daele, Volders, & Vlaeyen, 2016). It is also important to note that while learning theory may provide a parsimonious explanation of numerous key mechanisms in the etiology, maintenance, and treatment of avoidance and anxiety disorders, it may represent only a portion of the complexity inherent in psychopathology. However, given the centrality of learning theory to anxiety disorders, the present article attempted to provide a more nuanced analysis of avoidance behavior in order to determine translational treatment approaches that may be useful as adjuncts to current exposure-based procedures.

References

Baetu, I. & Baker, A. G. (2010). Extinction and blocking of conditioned inhibition in human causal learning. *Learning & Behavior, 38*, 394-407. doi:10.3758/LB.38.4.394

Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learning & Memory, 11*, 485-494. doi: 10.1101/lm.78804

Bravo-Rivera, C., Roman-Ortiz, C., Montesinos-Cartagena, M., and Quirk, G. J. (2015). Persistent active avoidance correlates with activity in prelimbic cortex and ventral striatum. *Front. Behav. Neurosci.* 9:184. doi: 10.3389/fnbeh.2015.00184

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

Davis, M., Ressler, K., Rothbaum, B. O., & Richardson, R. (2006). Effects of D-cycloserine on extinction: translation from preclinical to clinical work. *Biological psychiatry, 60*(4), 369-375. doi.org/10.1016/j.biopsych.2006.03.084

Declercq M., De Houwer J. (2008). Evidence for the interchangeability of an avoidance behavior and a negative occasion setter. *Learning & Behavior, 36*, 290–300. 10.3758/LB.36.4.290

Declercq, M., De Houwer, J., & Baeyens, F. (2008). Evidence for an expectancy-based theory of avoidance behaviour. *The Quarterly Journal of Experimental Psychology, 61*, 1803-1812. doi: 10.1080/17470210701851214

DeVito, P. L., & Fowler, H. (1987). Enhancement of conditioned inhibition via an extinction treatment. *Animal Learning & Behavior, 15*(4), 448-454. doi:10.3758/BF03205055

Foilb, A. R., Flyer-Adams, J. G., Maier, S. F., & Christianson, J. P. (2016). Posterior insular cortex is necessary for conditioned inhibition of fear. *Neurobiology of Learning and Memory, 134*, 317-327. doi.org/10.1016/j.nlm.2016.08.004

Ginsburg, G. S., Becker, E. M., Keeton, C. P., Sakolsky, D., Piacentini, J., Albano, A. M…Kendall, P. C. (2014). Naturalistic follow-up of youths treated for pediatric anxiety disorders. *JAMA Psychiatry, 71*, 310-318. doi:10.1001/jamapsychiatry.2013.4186

Guastella, A. J., Richardson, R., Lovibond, P. F., Rapee, R. M., Gaston, J. E., Mitchell, P., & Dadds, M. R. (2008). A randomized controlled trial of D-cycloserine enhancement of exposure therapy for social anxiety disorder. *Biological psychiatry, 63*(6), 544-549.

Haaker, J., Gaburro, S., Sah, A., Gartmann, N., Lonsdorf, T. B., Meier, K., ... & Kalisch, R. (2013). Single dose of L-dopa makes extinction memories context-independent and prevents the return of fear. *Proceedings of the National Academy of Sciences, 110*(26), E2428-E2436. doi/10.1073/pnas.1303061110

Hikind, N., & Maroun, M. (2008). Microinfusion of the D1 receptor antagonist, SCH23390 into the IL but not the BLA impairs consolidation of extinction of auditory fear conditioning. *Neurobiology of learning and memory, 90*(1), 217-222. doi.org/10.1016/j.nlm.2008.03.003

Harmer, C. J., & Phillips, G. D. (1999). Enhanced conditioned inhibition following repeated pretreatment with d-amphetamine. *Psychopharmacology, 142*(2), 120-131. doi:10.1007/s002130050870

Hermann, A., Stark, R., Milad, M. R., & Merz, C. J. (2016). Renewal of conditioned fear in a novel context is associated with hippocampal activation and connectivity. *Social cognitive and affective neuroscience*, nsw047. doi: 10.1093/scan/nsw047

Holland, P. C. (1984). Differential effects of reinforcement of an inhibitory feature after serial and simultaneous feature negative discrimination training. *Journal of experimental psychology. Animal behavior processes, 10*, 461-475.

Holland, P. C. (1986). Temporal determinants of occasion setting in feature-positive discriminations. *Animal Learning & Behavior, 14*(2), 111-120. doi:10.3758/BF03200045

Holland, P. C. (1989). Feature extinction enhances transfer of occasion setting. *Animal Learning & Behavior, 17*(3), 269-279. doi:10.3758/BF03209799

Holland, P. C., & Bouton, M. E. (1999). Hippocampus and context in classical conditioning. *Current opinion in neurobiology, 9*(2), 195-202. doi.org/10.1016/S0959-4388(99)80027-0

Holtzman-Assif, O., Laurent, V., & Westbrook, R. F. (2010). Blockade of dopamine activity in the nucleus accumbens impairs learning extinction of conditioned fear. *Learning & Memory, 17*(2), 71-75. doi: 10.1101/lm.1668310

Iordanova, M. D., McNally, G. P., & Westbrook, R. F. (2006). Opioid receptors in the nucleus accumbens regulate attentional learning in the blocking paradigm. *The*

Journal of neuroscience, *26*(15), 4036-4045. doi: 10.1523/JNEUROSCI.4679-05.2006

Krypotos, A., Effting, M., Kindt, M., and Beckers, T. (2015). Avoidance learning: a review of theoretical models and recent developments. *Front. Behav. Neurosci*. 9:189. doi: 10.3389/fnbeh.2015.00189

Li, S. S. Y., & McNally, G. P. (2014). The conditions that promote fear learning: prediction error and Pavlovian fear conditioning. *Neurobiology of learning and memory*, *108*, 14-21. doi.org/10.1016/j.nlm.2013.05.002

Laurent, V., Wong, F. L., & Balleine, B. W. (2015). δ- Opioid receptors in the accumbens shell mediate the influence of both excitatory and inhibitory predictions on choice. *British journal of pharmacology*, *172*(2), 562-570. DOI: 10.1111/bph.12731

Lissek, S., Glaubitz, B., Wolf, O. T., & Tegenthoff, M. (2015). The DA antagonist tiapride impairs context-related extinction learning in a novel context without affecting renewal. *Frontiers in behavioral neuroscience*, *9*. doi: 10.3389/fnbeh.2015.00238

Lissek, S., Levenson, J., Biggs, A. L., Johnson, L. L., Ameli, R., Pine, D. S., & Grillon, C. (2008). Elevated fear conditioning to socially relevant unconditioned stimuli in social anxiety disorder. *The American Journal of Psychiatry*, *165*, 124-132. doi.org/10.1176/appi.ajp.2007.06091513

Lovibond, P.F., Chen, S.X., Mitchell, C.J. & Weidemann, G. (2013). Competition between an avoidance response and a safety signal: Evidence for a single learning system. *Biological Psychology*, *92*, 9-16

Lovibond, P.F., Mitchell, C.J., Minard, E., Brady, A., Menzies, R.G. (2009). Safety behaviours preserve threat beliefs: Protection from extinction of human fear conditioning by an avoidance response. *Behaviour Research and Therapy 47*, pp. 716-720.

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review, 82*, 276-298.

Melchers, K. G., Lachnit, H., & Shanks, D. R. (2004). Within-compound associations in retrospective revaluation and in direct learning: A challenge for comparator theory. *Quarterly Journal of Experimental Psychology*, **57B**, 25–53.

Morell, J. R., & Holland, P. C. (1993). Summation and transfer of occasion setting. *Animal Learning & Behavior, 21*, 145-153.

Mowrer O. H. (1951). Two-factor learning theory: summary and comment. *Psychological Review, 58*, 350–354. doi.org/10.1037/h0058956

Mueller, D., Bravo-Rivera, C., & Quirk, G. J. (2010). Infralimbic D2 receptors are necessary for fear extinction and extinction-related tone responses. *Biological psychiatry, 68*(11), 1055-1060. doi.org/10.1016/j.biopsych.2010.08.014

Myers, K. M., & Davis, M. (2007). Mechanisms of fear extinction. *Molecular psychiatry, 12*(2), 120-150. doi:10.1038/sj.mp.4001939

Meyer, H. C., Putney, R. B., & Bucci, D. J. (2015). Inhibitory learning is modulated by nicotinic acetylcholine receptors. *Neuropharmacology, 89*, 360-367. doi.org/10.1016/j.neuropharm.2014.10.025

Meulders, A., Van Daele, T., Volders, S., & Vlaeyen, J. W. (2016). The use of safety-seeking behavior in exposure-based treatments for fear and anxiety: Benefit or

burden? A meta-analytic review. *Clinical psychology review*, *45*, 144-156. doi.org/10.1016/j.cpr.2016.02.002

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, *87*, 532-552.

Pearce, J. M., & Wilson, P. N. (1991). Failure of excitatory conditioning to extinguish the influence of a conditioned inhibitor. Journal of experimental *psychology. Animal Behavior Processes, 17,* 519-529.

Quirk, G. J., & Mueller, D. (2008). Neural mechanisms of extinction learning and retrieval. *Neuropsychopharmacology, 33,* 56-72. doi:10.1038/sj.npp.1301555

Rescorla, R. A. (1986). Extinction of facilitation. *Journal of Experimental Psychology: Animal Behavior Processes*, *12*(1), 16. doi.org/10.1037/0097-7403.12.1.16

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. A.H. Black, W.F. Prokasy (Eds.), Classical conditioning II: Current research and theory, Appleton-Century-Crofts, New York (1972), pp. 64–99

Rodriguez-Romaguera, J., Greenberg, B. D., Rasmussen, S. A., & Quirk, G. J. (2016). An avoidance-based rodent model of exposure with response prevention therapy for obsessive-compulsive disorder. *Biological Psychiatry*. doi.org/10.1016/j.biopsych.2016.02.012

Scheveneels, S., Boddez, Y., Vervliet, B., & Hermans, D. (2016). The validity of laboratory-based treatment research: Bridging the gap between fear extinction and exposure treatment. *Behaviour Research and Therapy*, *86*, 87-94.

Schultz, W. (2007). Behavioral dopamine signals. *Trends in neurosciences*, *30*(5), 203-210. doi.org/10.1016/j.tins.2007.03.007.

Shiflett, M. W., Riccie, M., & DiMatteo, R. (2013). The effects of amphetamine sensitization on conditioned inhibition during a Pavlovian–instrumental transfer task in rats. *Psychopharmacology*, *230*(1), 137-147. doi:10.1007/s00213-013-3144-3

Solomon, R. L., Kamin, L. J., and Wynne, L. C. (1953). Traumatic avoidance learning: the outcomes of several extinction procedures with dogs. *J. Abnorm. Soc. Psychol.* 48, 291–302. doi: 10.1037/h0058943

Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., & Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nature neuroscience*, *16*(7), 966-973. doi:10.1038/nn.3413

Vansteenwegen, D., Hermans, D., Vervliet, B., Francken, G., Beckers, T., Baeyens, F., & Eelen, P. (2005). Return of fear in a human differential conditioning paradigm caused by a return to the original acquisition context. *Behaviour Research and Therapy, 43*, 323-336. doi.org/10.1016/j.brat.2004.01.001

van Vooren, P. R., Franssen, M., Beckers, T., Hermans, D., & Baeyens, F. (2012). Narrowing down the conditions for extinction of Pavlovian feature-positive discriminations in humans. *Learning & behavior*, *40*(4), 393-404. doi:10.3758/s13420-011-0060-4

Vervliet, B., & Indekeu, E. (2015). Low-cost avoidance behaviors are resistant to fear extinction in humans. *Frontiers in Behavioral Neuroscience, 9*, 351. doi: 10.3389/fnbeh.2015.00351

Yoon, T., Graham, L. K., & Kim, J. J. (2011). Hippocampal lesion effects on occasion setting by contextual and discrete stimuli. *Neurobiology of learning and memory*, *95*(2), 176-184. doi.org/10.1016/j.nlm.2010.07.001

Yi, F., Catudio-Garrett, E., Gábriel, R., Wilhelm, M., Erdelyi, F., Szabo, G., ... & Lawrence, J. (2015). Hippocampal “cholinergic interneurons” visualized with the choline acetyltransferase promoter: anatomical distribution, intrinsic membrane properties, neurochemical characteristics, and capacity for cholinergic modulation. *Frontiers in synaptic neuroscience*, *7*, 4. doi.org/10.3389/fnsyn.2015.00004

Zelikowsky, M., Hast, T. A., Bennett, R. Z., Merjanian, M., Nocera, N. A., Ponnusamy, R., & Fanselow, M. S. (2013). Cholinergic blockade frees fear extinction from its contextual dependency. *Biological psychiatry*, *73*(4), 345-352. doi.org/10.1016/j.biopsych.2012.08.006

Zimmer-Hart, C. L., & Rescorla, R. A. (1974). Extinction of Pavlovian conditioned inhibition. *Journal of Comparative and Physiological Psychology, 86,* 837-845.

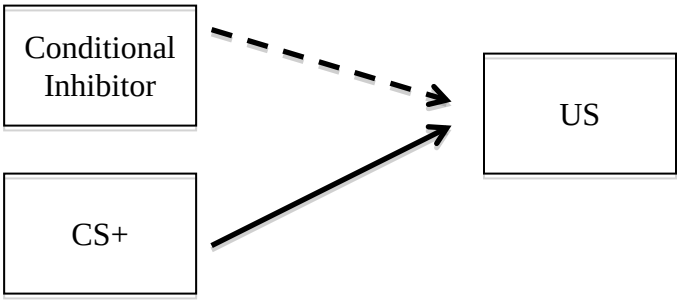

Figure 1: Relationship of a conditional inhibitor to a US. Dashed lines represent direct inhibitor associations, whereas solid lines represent direct excitatory associations.

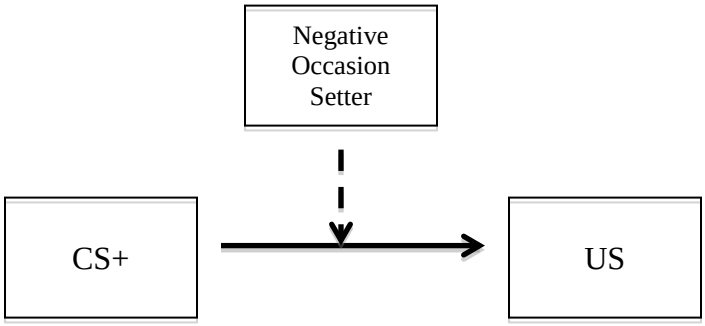

Figure 2: Modulatory impact of a negative occasion setter on CS/US relationship.