May 2020

Neural Substrates of Active Avoidance and Its Impact on Fear Extinction

Elizabeth Parisi
University of Wisconsin-Milwaukee

NEURAL SUBSTRATES OF ACTIVE AVOIDANCE AND ITS IMPACT ON FEAR EXTINCTION

by

Elizabeth A. Parisi

A Thesis Submitted in
Partial Fulfillment of the
Requirements for the Degree of

Master of Science
in Psychology

at
The University of Wisconsin-Milwaukee
May 2020

ABSTRACT

NEURAL SUBSTRATES OF ACTIVE AVOIDANCE AND ITS IMPACT ON FEAR EXTINCTION

by

Elizabeth A. Parisi

The University of Wisconsin-Milwaukee, 2020
Under the Supervision of Professor Christine L. Larson, Ph.D.

Models of anxiety suggest that avoidance of a conditioned fear stimulus prevents new safety learning, thereby serving to maintain fear. However, there is little empirical data in humans on the impact of avoidance of conditioned fear stimuli on subsequent fear extinction. In the present study I investigated the effect of avoidance of threat on neural activity during avoidance/control and a subsequent extinction phase using ultra high-resolution (7T) fMRI. Results indicated that active avoidance was associated with increased activity in regions involved in reward prediction, but this did not differentiate active avoidance from an active control condition. Neural activation during the extinction task appeared to support extinction learning and fear suppression in participants who previously engaged in active avoidance. These findings suggest that engagement in active avoidance did not impair new safety learning.

TABLE OF CONTENTS

PAGE

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Christine Larson, and the rest of my thesis committee, Drs. Hanjoo Lee and Krista Lisdahl. This thesis would not have been possible without your guidance and support. I would also like to thank Dr. Greg Hajcak for his support and collaboration on this project. Finally, I would like to thank my dear colleagues in the Affective Neuroscience Lab – Kenneth Bennett, Ashley Huggins, Carissa Weis, and Kate Webb – for their help with data collection and support during the thesis writing process.

Anxiety disorders are the most prevalent class of mental illness in the U.S., are relatively stable over time, and are complicated by high rates of comorbidity with other mood and anxiety psychopathology (Kessler et al., 2005; Kessler, Ruscio, Shear, & Wittchen, 2009). The presence of an anxiety disorder is predictive of disability, poor functional outcomes, decreased productivity, higher utilization of healthcare resources and increased morbidity and mortality (Stein et al., 2005; Bystritsky, Khalsa, Cameron, & Schiffman, 2013). Thus, anxiety-related disorders represent a significant public health burden (Kessler et al., 2005). On an individual level, anxiety disorders are associated with immense subjective suffering and reduced perceived quality of life (Barrera & Norton, 2009). Moreover, anxiety disorders alone have been shown to be significantly associated with suicidal ideation and suicide attempts cross-sectionally and longitudinally, with comorbid anxiety and mood disorders conferring greater risk for suicide attempts than mood disorders alone (Norton, Temple, & Pettit, 2008; Sareen et al., 2005). Although several empirically-supported treatments exist for anxiety disorders, many patients fail to benefit fully from treatment or relapse. Consequently, investigations into the neural substrates of maladaptive anxiety and the mechanisms through which it is maintained are crucial in the optimization of treatment of clinical anxiety (Brooks & Stein, 2015).

Threat avoidance plays a role in the development and maintenance of clinical anxiety

There is significant empirical support for conditioning models of anxiety disorders, which implicate dysregulation of fear processing in their pathogenesis (Lissek et al., 2005; Makinson & Young, 2012). Specifically, individuals with pathological anxiety show stronger acquisition of conditioned fear (suggesting heightened excitatory fear processes), and weaker extinction of conditioned fear (suggesting decreased inhibitory fear processes) (Lissek et al., 2005). Since the genesis of the conditioning model of anxiety disorders, the theory has expanded to include

negative reinforcement of cognitive and behavioral avoidance as a mechanism contributing to their development and maintenance (Lissek et al., 2005). While avoidance of threat is generally adaptive and serves to protect an organism from real, imminent danger, this process can become maladaptive when contingencies do not reflect reality or fear responses are excessive or inappropriate (Makinson & Young, 2012). Excessive attempts to avoid thoughts, feelings or external stimuli associated with negative emotional states are prominent and maladaptive features of anxiety disorders and are thought to maintain anxiety by becoming a chronic strategy for coping with distressing thoughts and emotions (American Psychiatric Association, 2013; Schlund & Cataldo, 2010a; Schlund et al., 2010b). Individual difference characteristics associated with increased risk for pathological anxiety, such as trait anxiety, intolerance of uncertainty, and behavioral inhibition, have been associated with increased rate and duration of behavioral avoidance of threat (Carleton et al., 2012; Sheynin et al., 2014; Spielberger, Sydeman, Owen, & Marsh, 1999). Across disorders, avoidance is associated with increased fear and catastrophic thoughts and serves to enhance and maintain anxiety and physiological reactions over the long-term (Schlund et al., 2010b). Exposure therapy, based on the principles of extinction, has received substantial empirical support for treatment of anxiety disorders. However, avoidance behaviors are thought to interfere with success of exposure therapy, which aims to reduce fear through disconfirmation of excessive threat beliefs, by blocking opportunities for extinction and impairing extinction learning (Blakey & Abramowitz, 2016; Lovibond, Mitchell, Minard, Brady, & Menzies, 2009; Lovibond, Chen, Mitchell, & Weidemann, 2013). Therefore, understanding the impact of avoidance on fear extinction behaviorally and neurologically may offer insight into factors that hinder the efficacy of exposure therapy for anxiety disorders.

Acquisition and extinction of conditioned fear in humans

Pavlovian fear conditioning is the process by which a neutral stimulus is paired with an aversive stimulus (unconditioned stimulus, US). Following repeated pairing with the US, the neutral stimulus becomes aversive itself (conditioned stimulus, CS), signaling onset of the US and evoking a fearful response (conditioned response, CR) in anticipation of the US (Lissek et al., 2005). In extinction, the acquired fearful response to the CS decreases when it is no longer reinforced by association with the aversive US (Quirk & Mueller, 2008). In other words, while fear conditioning is the process by which a previously safe stimulus is associated with threat, extinction is the process by which a previously threatening stimulus is established as safe (Myers & Davis, 2007). Extinction learning can be studied experimentally using neural and psychophysiological indices of fear reactivity, with decreased frequency and magnitude (i.e., suppression) of the conditioned response indicating successful extinction (McNally, 2007). It is important to note that extinction of conditioned fear does not erase previous CS-US associations, but rather represents the development of inhibitory associations that compete with previously learned ones (McNally, 2007; Quirk, Garcia, & González-Lima, 2006). Spontaneous recovery, a phenomenon in which a CR reemerges naturally following extinction, supports the idea that extinction does not abolish CS-US associations but rather suppresses them (McNally, 2007). Thus, deficits in fear extinction are thought to represent a deficit in the suppression of CS-US associations (Rauch, Shin, & Phelps, 2006).

Avoidance behavior interferes with extinction learning

"Protection from extinction" is a phenomenon that occurs when an inhibitory CS (i.e., a safety signal) is presented concurrently with an excitatory CS (i.e., conditioned threat signal) during extinction, such that each trial is perceived as safe due to the presence of the inhibitory

CS and extinction of the fearful response to the excitatory CS cannot occur (Lovibond, Davis, & O'Flaherty, 2000). "Protection from extinction" has been demonstrated experimentally through the introduction of an inhibitory CS/voluntary safety behavior to avoid an aversive stimulus (Lovibond et al., 2009). The opportunity to engage in a voluntary avoidance behavior (e.g., pressing a button) during presentation of an excitatory CS has been shown to diminish fear responding (i.e., skin conductance, threat expectancy ratings). Fear responses remain attenuated during extinction while the avoidance behavior is available, with a sharp rebound of physiological and subjective indices of fear when the avoidance behavior is eliminated (Lovibond et al., 2009). It has been shown that although extinction of fear towards the excitatory CS is possible following removal of an avoidance behavior, restoration of the avoidance behavior following extinction tends to occur (indicating spontaneous recovery and resistance of avoidance behaviors to fear extinction) (Vervliet & Indekeu, 2015). Spontaneous recovery of conditioned fear and reestablishment of avoidance behavior has been shown to be more robust in trait anxious individuals, suggesting that this effect is particularly strong in individuals at risk for pathological anxiety (Vervliet & Indekeu, 2015). This provides support for the idea that avoidance behavior cancels the threat expectancy generated by a CS, thereby inhibiting extinction learning. There is a dearth of empirical evidence characterizing the impact of avoidance on neural substrates of extinction in humans. Elucidating the neural mechanisms underlying the phenomenon of protection from extinction will further explicate the role of within-situation safety behaviors in preserving threat beliefs and anxiety in patients undergoing exposure therapy.

Neural correlates of acquisition and extinction of conditioned fear

Amygdala. The amygdala is a key component of mammalian fear processing, mediating survival functions through the coordination of defensive responding to potentially threatening stimuli using both interoceptive and exteroceptive cues (Makinson & Young, 2012; LaBar, Gatenby, Gore, LeDoux, & Phelps, 1998). Considerable evidence from neuroimaging studies in humans suggest that the amygdala plays a crucial role in the generation, expression and experience of negative emotional reactions and physiological reactivity in response to potentially threatening stimuli (Bryant et al., 2008; Makinson & Young, 2012). Several neuroimaging studies have demonstrated an association between activity changes in the amygdala and changes in peripheral indices of fear responding (e.g., skin conductance) and correlations between activity in the amygdala and thalamus, which projects to the HPA axis to initiate physiological stress responding (Rauch et al., 2006). Neuroimaging studies have also observed increased activation in the amygdala during cued conditioning, suggesting that it plays a role in the acquisition of conditioned fear (Rauch et al., 2006). Indeed, selective bilateral damage to the amygdala has been shown to impair the acquisition of fear, such that the individual can explicitly state which stimulus was paired with a US but does not display implicit changes in physiological reactivity that would suggest a CR (Makinson & Young, 2012; McNally, 2007). Via its connections to the prefrontal cortex (PFC), the amygdala also plays a role in the retention of emotional memories (Makinson & Young, 2012). Additionally, greater amygdala activation in response to the CS- compared to the CS+ predicts increased success of extinction, indicating that the amygdala may play a role in early extinction learning (Rauch et al., 2006). Considered part of the extended amygdala – a cluster of highly structurally-connected brain regions with similar ontogeny, cytoarchitecture and functions – the bed nucleus of the stria terminalis (BNST) has

been implicated in the expression of fear, with BNST lesions resulting in attenuated fear responses to contextual stimuli (Sullivan et al., 2004). Additionally, the BNST plays a role in the anticipation of aversive events, particularly in uncertain contexts (Avery, Clauss, & Blackford, 2016; Avery et al., 2014). However, investigation of the role of the BNST using neuroimaging in humans is limited due to its very small size in the human medial basal forebrain and the relatively low resolution of functional magnetic resonance imaging (fMRI) (Avery et al., 2016).

Medial prefrontal and anterior cingulate cortices. In humans, substantial evidence from neuroimaging investigations has implicated the medial prefrontal (mPFC) and anterior cingulate (ACC) cortices in emotional processing. Specifically, ventral-rostral regions of the ACC/mPFC play a regulatory role in the expression of emotional responses through inhibitory projections to subcortical structures (e.g., amygdala) while dorsal-caudal regions are involved in the expression and appraisal of negative emotion (Etkin, Egner, & Kalisch, 2011). Similarly, the ACC/mPFC play a role in the suppression of conditioned fear (i.e., fear extinction) through inhibitory projections to the amygdala (Felmingham et al., 2007; McNally, 2007; Etkin et al., 2011). While lesions of the mPFC in rats do not impact the acquisition of conditioned fear, they have been shown to significantly impair extinction recall, such that rats with lesions of the mPFC show difficulty extinguishing fear over multiple sessions and show impaired memory for extinction following a delay (Quirk et al., 2006). In humans, increased activation in the vmPFC and ACC and decreased activation in the amygdala is observed during successful fear extinction and extinction recall (Quirk et al., 2006; Etkin et al., 2011). Following successful extinction, mPFC activation in response to the CS+ increases, indicating that mPFC activity is potentiated by extinction (Quirk et al., 2006). This suggests that the mPFC comes online during extinction to excite inhibitory pathways responsible for reducing the expression of fear (Quirk et al., 2006).

Indeed, electrical stimulation of the mPFC results in decreased activity in projections from the central nucleus of the amygdala to the brainstem, thereby decreasing the expression of conditioned fear (Quirk et al., 2006). Further, increased mPFC activity is associated with increased extinction behavior (Quirk et al., 2006). Electrical stimulation and metabolic enhancement of the mPFC has been shown to strengthen extinction memories, such that short-term extinction memory is unaffected but retention of extinction (evidenced by decreases in spontaneous recovery over a delay) is markedly improved (Quirk et al., 2006).

Hippocampus. Fewer neuroimaging studies investigating the role of the hippocampus in human fear processing have been published to date. In the animal literature, there is substantial evidence to suggest that the hippocampus is not required for the acquisition of cued fear (as indicated by lesion studies) but it seems to play an important role in the acquisition of aversive context conditioning (Brooks & Stein, 2015; Marschner, Kalisch, Vervliet, Vansteenwegen, & Büchel, 2008). Studies using pharmacological inactivation of the hippocampus indicate that inactivation of the hippocampus prior to extinction training leads to poor subsequent recall of extinction (Corcoran, Desmond, Frey, & Maren, 2005; Quirk & Mueller, 2008). This suggests that hippocampal activation and plasticity is necessary for recall of extinction within the conditioned context (Quirk & Mueller, 2008). Similar findings in the mPFC suggest that the mPFC and hippocampus may interact for contextual modulation of extinction recall, which is crucial for accurately distinguishing contextual cues that indicate safety versus those that indicate threat (Quirk & Mueller, 2008; Sotres-Bayon, Cain, & LeDoux, 2006). These findings translate to neuroimaging findings in humans, such that a human lesion study demonstrated a double dissociation between patients with bilateral damage to either the amygdala or the hippocampus. It was observed that individuals with hippocampal damage failed to demonstrate declarative

knowledge of the conditioning (i.e., which stimulus was the threat cue) but acquired an implicit conditioned response to the threat cue, while the opposite pattern was observed in those with amygdala damage (Makinson & Young, 2012; McNally, 2007). This pattern of results and findings from other neuroimaging investigations suggest that while the amygdala is necessary for the acquisition of a fear response to the CS, the hippocampus is necessary for acquisition of information related to the context in which conditioning occurred and context-US associations (Bechara et al., 1995; Marschner et al., 2008).

Neural correlates of threat avoidance

Although avoidance behavior has been implicated in the pathogenesis of many clinical disorders, progress has been limited in understanding the neurocircuitry supporting active avoidance of threat in humans (Schlund et al., 2010b). Imbalances in neural substrates that process reward-motivated approach behavior and aversively-motivated avoidance behavior have been implicated in the development of dysfunctional avoidance coping (Schlund, Magee, & Hudgins, 2011). Avoidance is motivated by threatening cues and leads to withdrawal from threat. This resultant removal of threat, and consequent fear reduction, negatively reinforce avoidance leading to a chronic pattern of avoidance coping (Schlund et al, 2010b). Neuroimaging in humans provides evidence indicating significant overlap between neural systems contributing to approach and avoidance behavior, suggesting that avoidance recruits neurocircuitry associated with reward processing and may be intrinsically rewarding (Schlund et al., 2010b; Schlund et al., 2011). A distributed fronto-limbic-striatal network is implicated in reward learning, including regions involved in cognitive/behavioral regulation (anterior cingulate, superior and medial frontal regions), emotional cue salience processing (insula, amygdala) and establishing response-outcome relations (striatum) (Schlund et al., 2011). Avoidance cues and outcomes have been

shown to activate the amygdala, insula, striatum and medial frontal regions in humans, supportive of the role of avoidance as a negative reinforcer (Schlund et al., 2011).

Reward neurocircuitry recruited in avoidance. The ventral striatum, which encompasses the nucleus accumbens (NAcc), ventral caudate and ventral putamen, plays a crucial role in the processing of appetitive stimuli, with increased activation and dopamine release in the ventral striatum evident during anticipation of pleasurable events (Jensen et al., 2003). However, human neuroimaging studies have also demonstrated consistent increases in ventral striatum activation in anticipation of highly-salient aversive events (e.g., an electric shock) (Jensen et al., 2003; Pohlack, Nees, Ruttorf, Schad, & Flor, 2012). Further, the ventral striatum has been implicated in the processing of threat cues and coordination of escape behavior in avoidance of shock, pain, or aversive images or sounds (Bolstad et al., 2013; Delgado, Nystrom, Fissell, Noll, & Fiez, 2000; Eldar, Hauser, Dayan, & Dolan, 2016; Jensen et al., 2003; Levita, Hoskin, & Champi, 2012; Schlund et al., 2016; Seymour, Daw, Dayan, Singer, & Dolan, 2007). Appropriately, the ventral striatum receives input from structures involved in motivational processes (i.e., insula, hippocampus, amygdala, prefrontal cortex) and primarily projects to the ventral global padillus, which is involved in the regulation of voluntary movement. Thus, via its connections to frontal and limbic sites, the ventral striatum is crucial in the coordination of both appetitively- and aversively-motivated behavior, including threat avoidance (Jensen et al., 2003). While the habenula has not been well-investigated in humans due to its small size, animal studies suggest that it is crucially involved in modulation of motivated behavior including reward prediction and behavioral avoidance (Namboodiri, Rodriguez-Romaguera, & Stuber, 2016). Specifically, lateral habenula (LHb) activation during loss of reward is associated with inhibition of dopaminergic neurons, as well as deficits in the acquisition of avoidance behavior. Accordingly, LHb lesions

have been shown to accelerate avoidance learning suggesting that the LHb is involved in inhibition of dopaminergic activity underlying reinforcement of avoidance behaviors. While stimulation of the LHb has been shown to inhibit acquisition of avoidance behaviors, stimulation of the ventral tegmental area (VTA) has been shown to rapidly increase the rate of avoidance acquisition. It is suggested that the LHb and VTA act as mutually inhibitory, such that stimulation of the LHb is associated with substantial decreases in dopaminergic activity in VTA neurons (and thus slower avoidance acquisition), and increased activity of dopaminergic neurons in the VTA inhibits LHb activity (Shumake, Ilango, Scheich, Wetzel, & Ohl, 2010).

Threat neurocircuitry implicated in avoidance. Animal studies have demonstrated that the amygdala plays a crucial role in avoidance by signaling cues that predict delivery of an aversive stimulus. In humans, the role of the amygdala in avoidance is less clear. While avoidance is consistently associated with activation in the striatum, the amygdala consists of subnuclei that have been implicated in different aspects of avoidance learning. A small subset of human neuroimaging studies have investigated differentiation in the roles of amygdala subnuclei in avoidance, particularly the basolateral amygdala (BLA) and central nucleus of the amygdala (CeA) (Schlund et al., 2010b). The BLA shows increased activation to threatening avoidance cues relative to baseline, suggesting that it plays a role in avoidance learning. It has also been shown to be involved in recall of avoidance learning, such that lesions of the BLA have been shown to impair the acquisition and recall of active avoidance, while lesions of the CeA do not impair avoidance recall but do inhibit acquisition (Ilango et al., 2014a; Ilango, Shumake, Wetzel, & Ohl, 2014b). Interaction between the BLA and the striatum has been associated with the learning of active avoidance as a means of temporarily weakening a conditioned response to feared stimuli (Delgado et al. 2000). Correspondingly, the BNST shares anatomical and

functional connections with the amygdala and striatum (Avery et al., 2016). While rodent studies provide evidence suggesting that the BNST plays a central role in sustained threat monitoring, few studies have investigated the role of the human BNST in fear and avoidance learning (Avery et al., 2016; Davis, 1998; Davis, Walker, Miles, & Grillon, 2010; Lebow & Chen, 2016; Shackman & Fox, 2016; Tyszka & Pauli, 2016).

A limitation of previous human neuroimaging studies is the use of 1.5T and 3T fMRI, which offers relatively lower spatial resolution and poses challenges for investigating the role of small structures such as the amygdala subnuclei, BNST, habenula and substructures of the ventral striatum. Using a 7T scanner, we can obtain higher-resolution images of these structures in order to further differentiate their roles in avoidance learning (Schlund et al., 2010b).

Preliminary investigation into the neural substrates of avoidance and extinction in humans

Although avoidance behavior is recognized as a core feature of many clinical disorders and is thought to interfere with the success of exposure-based therapies, there is a dearth of neurobiological evidence to suggest that avoidance impairs extinction learning (Schlund et al., 2010b). In one study investigating the impact of avoidance on neural and psychophysiological indices of extinction (Boeke, Moscarello, LeDoux, Phelps, & Hartley, 2017), subjects underwent fear acquisition followed by either an active avoidance or yoked extinction condition without active avoidance. In a subsequent session, subjects underwent extinction retrieval and novel acquisition. Results indicated significant between-group differences in conditioned responding (indexed by skin conductance) during the extinction retrieval phase, such that subjects who underwent yoked extinction showed an increase in conditioned responding from the yoked extinction phase to the extinction retrieval phase, while subjects in the active avoidance group showed no change in response. This suggests that fear reduction was more effective and long-

lasting in subjects who underwent avoidance before extinction learning. Between-group differences in BOLD activity were observed during the avoidance/yoked extinction phase, such that subjects in the avoidance condition showed greater striatal activation, and to a modest extent during the late avoidance phase, greater vmPFC activation. Additionally, increased striatal activation was observed during presentation of the CS+ versus the CS- in both groups. Consistent with previous findings, increased activity in the striatum was observed on trials when no shock was delivered in the avoidance condition only, suggesting that the striatum is preferentially activated during active avoidance versus passive extinction. Within-group differences in the putamen, caudate and mPFC during presentation of the CS+ versus the CS- in subjects in the avoidance condition were also evident, suggesting suppression of the conditioned response. The conclusion drawn from these findings is that control over aversive stimuli may have the potential to promote resilience to previously learned conditioned responses or acquisition of conditioned responses to novel threat. Additionally, the suggestion is made that active avoidance may be more effective than extinction in reducing fear responses and may represent a better approach for treatment of anxiety disorders than exposure-based therapies.

Overall, these findings are surprising in that they do not correspond with clinical observations or past empirical investigations of the role of avoidance in anxiety-related disorders. Further, other empirical evidence in humans suggests that avoidance or safety behavior during exposure therapy impairs the reduction of fear (Blakey & Abramowitz, 2016; Lovibond et al., 2009; Volders, Meulders, De Peuter, Vervliet, & Vlaeyen, 2012). There are some notable methodological limitations of this study to consider. First, the conditioned stimulus was reinforced by a shock in only 40% of trials during fear acquisition. Intermittent CS-US pairing during conditioning has been shown to produce slower learning rates, weaken acquisition

and expression of the conditioned response, and delay extinction learning compared to continuous reinforcement (Dunsmoor, Bandettini, & Knight, 2007; Grady, Bowen, Hyde, Totsch, & Knight, 2016). Additionally, both the CS+ and CS- were negatively-valenced stimuli (fearful faces). Viewing of fearful faces has been shown to increase indices of autonomic reactivity (e.g., skin conductance) as well as activation in the amygdala and other threat processing regions, even when masked (Carlson, Cha, & Mujica-Parodi, 2013; Hariri, Tessitore, Mattay, Fera, & Weinberger, 2002; Johnstone et al., 2005; Öhman, Esteves, & Soares, 1995; Öhman, 2002; Pessoa, Japee, & Ungerleider, 2005; Whalen et al., 2004). Further, it has been demonstrated that fearful faces are highly salient threat stimuli, eliciting greater physiological arousal and more robust amygdala response relative to other threatening images (e.g., violent scenes) (Hariri et al., 2002). This evidence that fearful faces are processed as highly salient threat stimuli in an automatic and preattentive manner suggests a potential confound in the ability to detect differences in threat reactivity to the CS+ versus CS- in both groups. Moreover, there were considerable differences between motor and cognitive demands in the avoidance vs. control conditions, such that participants in the avoidance condition were instructed to learn how to move a dot within a matrix to prevent shocks while control subjects were instructed to make multiple button presses to match the average motor response executed by avoidance subjects. Finally, some participants received shocks in the early part of the extinction phase, which may have accounted for incomplete reextinction in the yoked extinction group and increased conditioned responding during extinction retrieval. The authors acknowledge the additional limitation of using standard resolution 3T fMRI, such that they were unable to investigate the role of amygdala subnuclei and the BNST in avoidance. Further studies are needed to clarify these discrepancies in findings (Boeke et al., 2017).

Current study

The current study aims to investigate the neurocircuitry underlying active avoidance of a conditioned threat stimulus, and the impact of avoidance on subsequent extinction learning. To that end, participants were examined in a 7T fMRI scan, during which they underwent fear acquisition followed by either an avoidance or extinction condition, followed by extinction retrieval. Participants also completed a series of questionnaires assessing traits and behaviors associated with risk for anxiety.

Aim 1: Characterize the neurocircuitry recruited during active avoidance of threat.

Hypothesis: Differences in activation of neurocircuitry associated with threat processing (amygdala, CeA, BLA, BNST) and reward processing (nucleus accumbens, putamen, caudate, habenula) will be observed during avoidance vs. non-avoidance.

Aim 2: Analyze the impact of avoidance on neural activation during extinction.

Hypothesis: Based on previous findings (Lovibond et al., 2009; Lovibond et al., 2013), between-group differences are expected such that subjects in the avoidance group will show impaired extinction, evidenced by increased activity in regions associated with threat responding (amygdala subnuclei, BNST) during the extinction condition relative to subjects in the extinction group.

Aim 3: Investigate the association between trait anxiety and success of extinction learning following behavioral threat avoidance vs. non-avoidance.

Hypothesis: Greater trait anxiety will be associated with greater impairment of extinction following avoidance.

Method

Participants

Data from 53 participants were collected at the Medical College of Wisconsin. Participants were undergraduate students from the University of Wisconsin-Milwaukee recruited via the UWM Psychology Department's research subject pool. Eligibility requirements included age between 18 and 55 years, right-handed and English-speaking. Exclusion criteria included contraindications to MRI (e.g., irremovable metal in the body, pregnancy, claustrophobia), use of certain medications (antipsychotics, anticonvulsants, mood stabilizers), a history of head trauma, neurological conditions (e.g., epilepsy), psychosis, or bipolar disorder.

Procedure

Participants completed an online prescreen to assess for initial eligibility and received a code to participate in the study. Participants who passed the prescreen were contacted by study personnel to complete an MRI safety screening. Participants provided written informed consent. All study sessions took place in the Daniel M. Soref Imaging Research Facility on the Medical College of Wisconsin campus and included a series of functional and structural MRI scans, blood draws, and a battery of self-report questionnaires. Participants were compensated with course credit and cash payment. All study procedures were approved by the University of Wisconsin-Milwaukee and Medical College of Wisconsin Institutional Review Boards.

Shock Work-Up

Prior to completing the avoidance task, a shock work-up was completed to determine the level of electrical stimulation (i.e., shock) that would be used for the duration of the task based on subjective ratings. Shocks were delivered using a Psychlab stimulator (Contact Precision Instruments, Cambridge, MA). Two electrodes were placed approximately two inches above the

participant's left ankle. Starting at a low level of electrical stimulation (~.6mA, duration=500ms), a series of shocks were delivered. After each individual shock, participants were asked to make a 0 to 10 rating (0 = "didn't feel anything", 10 = "painful, but tolerable"). Participants were informed that the level set should be "painful, but tolerable" and the level selected would be used for the entirety of the task.

Avoidance task

The avoidance task consisted of three phases: acquisition, Active Avoidance (AA)/Active Control (AC), and extinction. During acquisition, participants were conditioned to the threat (CS+) and safety (CS-) cues. The acquisition phase consisted of a total of 16 trials (8 CS+, 8 CS-) in which the participant was presented with either a circle or a square (*Figure 1a*). The stimulus established as the CS+ was 100% reinforced via co-termination with the shock. Stimuli were counterbalanced such that for half of the participants, the circle was the CS+, while for the other half the square was the CS+. Stimulus presentation was presented in a pseudorandomized order and stimuli appeared on the screen for 6000ms. Participants viewed a fixation during inter-trial intervals for 5000 to 9000ms (average duration 7000ms). Following acquisition, participants were assessed for explicit learning of the CS-US contingency by being asked to indicate by button press which stimulus, presented side-by-side, predicted the shock.

Participants were randomly assigned to either the AA or AC condition (*Figure 1b*). In the AA condition, participants were instructed that the shock could be avoided by executing a button press when a border appears around the shape (last 1000ms of stimulus presentation). In the AC condition, participants were also instructed to execute a button press when the border appeared but were not instructed that the shock could be avoided. To control for motor responses, participants were asked to execute a button press when the border appeared on both CS+ and CS-

trials. In both conditions, participants were presented with 8 CS+ and 8 CS- trials with 6000ms duration. Stimuli were presented in a pseudorandom order and ITI varied from 5000ms to 9000ms (average duration 7000ms).

Following either the AA/AC condition, all participants underwent an extinction phase in which they were presented with 8 CS+ and 8 CS- trials without delivery of the shock (*Figure 1c*). During this phase, participants were instructed to discontinue button presses. AA participants were informed that they would no longer be able to avoid the shock. After each phase, participants were asked to make subjective anxiety ratings for the CS+, CS- and overall block.

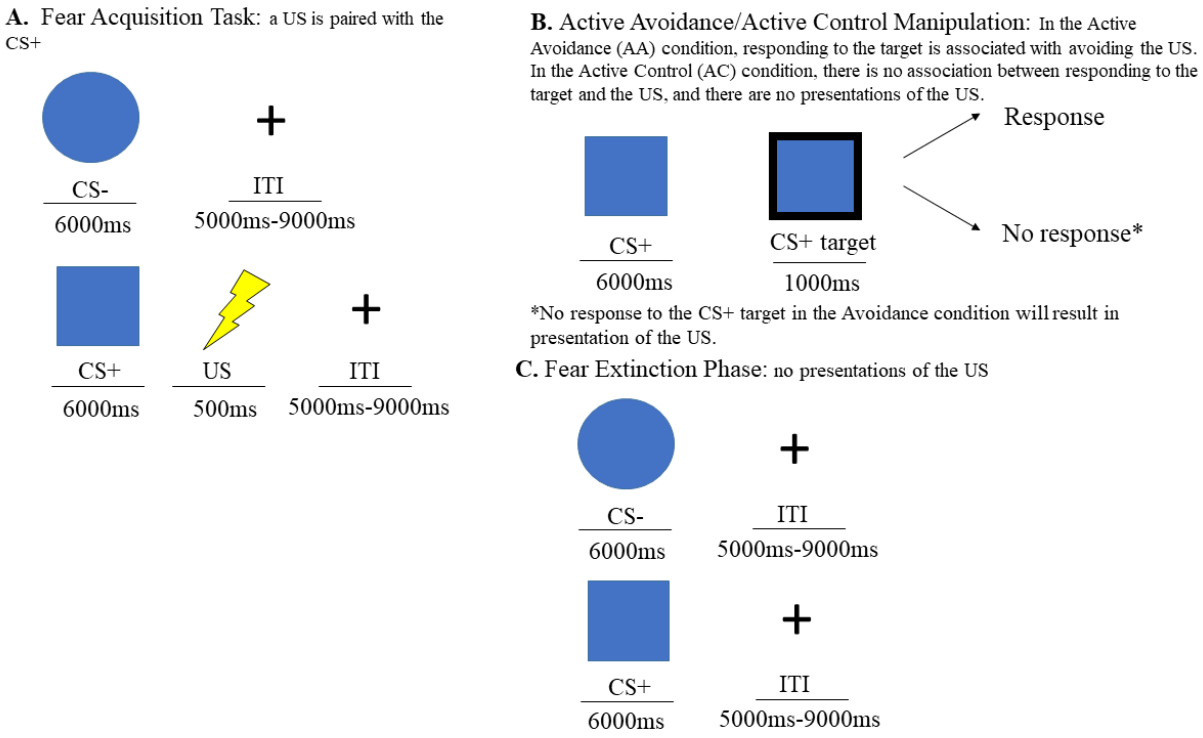

Figure 1. Avoidance task design. During acquisition (A), participants were presented with 8 trials each of CS+ (co-terminated with shock on 100% of trials) and CS-. During AA/AC (B), participants were presented with 8 trials each of the CS+ and CS-. The AA group was instructed to respond in order to avoid the shock while the AC group was instructed to simply respond. During extinction (C), participants were presented with 8 trials each of the CS+ (unreinforced) and CS-.

Trait Anxiety

Trait anxiety was measured using the Trait version of the Spielberger State-Trait Anxiety Inventory (STAI; Spielberger, 1983). The STAI consists of 20 self-report items rated on a four-point scale. The STAI has demonstrated good psychometric properties, including high test-retest reliability and internal consistency (Barnes, Harp, & Jung, 2002).

fMRI data acquisition

Anatomical. Imaging data were collected on a 7.0 Tesla MR950 General Electric scanner (GE Healthcare, Waukesha, WI). Whole-brain high-resolution T1-weighted anatomical images were acquired using a BRAVO gradient echo sequence with the following parameters: TR/TE = 8.012/3.784s; FOV: 220; flip angle = 5°; thickness = .8mm; matrix = 276 x 276; voxel size = 0.43 x 0.43 x 0.80mm. A high-resolution T2-weighted structural scan with partial coverage was collected in order to create anatomical ROIs. For the T2 structural scan, oblique images were acquired coronally, angulated perpendicular to the long axis of the hippocampal formation: TR/TE = 10000/30.66; FOV: 85; voxel size = 0.4297 x 0.4297 x 2mm.

Functional. Partial-brain functional T2*-weighted EPI scans were acquired in an axial orientation with the following parameters: TR/TE = 2500ms/24ms; flip = 73°; FOV = 220; matrix = 224 x 224; thickness = 1.8mm; voxel size = 0.8594 x 08594 x 1.8mm. Partial-brain coverage was optimized to take advantage of the high resolution capabilities of the 7T scanner and prioritize a priori ROIs of the study aims, including the amygdala (CeA, LA, BLA), BNST, striatum (caudate, nucleus accumbens, putamen), and habenula. Scan coverage was determined on an individual subject basis by placing the most inferior slice to cover the most ventral part of the hippocampus (*Figure 2*). An additional single-volume EPI scan with reverse phase encode

polarity was collected after the task to correct for susceptibility-related distortion during image processing.

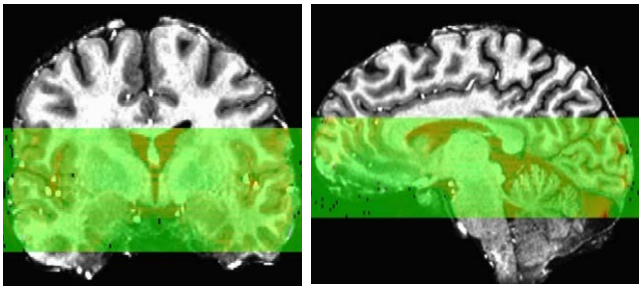

Figure 2. Example EPI partial coverage from a representative subject.

Preprocessing

Data was analyzed using Analysis of Functional Neural Images (AFNI) software (Cox, 1996). The first three volumes were removed to allow for scanner equilibration, and volumes with excessive motion (>2mm) and/or outliers (>10% of voxels in the volume identified as outliers) were censored. Due to greater sensitivity to distortion at ultra-high field, remaining EPI volumes were distortion corrected by warping to a middle space with the reverse phase encode polarity scan. EPI volumes were co-registered to the first functional volume, aligned to the subject's anatomy, and converted to percent signal change. A blur of 4mm FWHM was applied to the data. For whole brain group analyses, data was normalized to template (MNI152). Single subject BOLD responses were modeled with regressors for each condition type and relevant difference score (acquisition: CS+, CS-, CS+-CS-; AA/AC: CS+, CS-, CS+-CS-; extinction: CS+, CS-, CS+-CS-) for each voxel in the functional dataset. Due to the relative transience of neural responses to discrete cues and known habituation effects in fear acquisition and extinction, each task (acquisition, AA/AC, extinction) was divided into "early" and "late" phases (LaBar, Gatenby, Gore, LeDoux & Phelps, 1998). As each task (acquisition, AA/AC, extinction)

consisted of 8 trials total, “early-phase” included the first four trials, and “late-phase” the last four trials, of each task. Motion parameters were included as regressors of no interest.

Group Analyses

Univariate analyses. Two approaches were used for group analyses: voxelwise and region of interest (ROI) analysis. For the voxelwise analysis, images were converted to MNI 152 space using affine and nonlinear transformation. Voxelwise analysis of contrasts were identified using a $p < .001$ (uncorrected) and cluster probability of $p < .05$. A series of independent samples t-tests were conducted in AFNI 3dttest++ to compare strength of neural activation to the CS+ and differentiation between the CS+ and CS- (quantified as a difference score: CS+-CS-). To identify brain regions preferentially involved in avoidance (Aim 1), independent samples t-tests were conducted to examine between-subject differences in BOLD activity (CS+, CS+- CS-) in the AA and AC groups during early- and late-phase AA/AC task. To examine the effect of avoidance on neural activation during extinction (Aim 2), independent samples t-tests were conducted to examine between-subject differences in BOLD activity (CS+, CS+-CS-) in the AA and AC groups during early- and late-phase extinction. To examine conditioning effects in the fear acquisition task, a paired sample t-test comparing within-subject differences in BOLD activity to the CS+ versus CS- collapsed across AA and AC groups in early- and late-phase acquisition was conducted. To ensure group equivalency at baseline, an independent samples t-test comparing between-subject (AA versus AC) differences in discrimination between the CS+ and CS- (quantified as CS+-CS-) during early- and late-phase acquisition was conducted. Statistical thresholds for all tests were set at $\alpha = .01$.

A separate ROI analysis was conducted for amygdala subnuclei (LA, CeA, BLA), BNST, striatum (caudate, putamen, NAcc) and habenula based on a priori hypotheses. ROI

segmentation masks were defined by MNI standard space coordinates for the amygdala subnuclei (CeA, BLA, LA), striatal structures (caudate, putamen, nucleus accumbens) and habenula. The BNST ROI was defined by the segmentation mask constructed by Theiss and colleagues (2017). Mean beta coefficients for each ROI for each condition (CS+/CS-) for each subject was calculated using AFNI 3dROIstats. These extracted mean ROI data were used in a series of t-tests and 2x2 ANOVAs. To address Aim 1, mean beta coefficients for each condition (CS+, CS-) in all a priori ROIs in the AA/AC tasks were entered into a Group (AA, AC) by Condition (CS+, CS-) repeated-measures ANOVA. Also in the AA/AC task, differentiation between the CS+ and CS- (quantified as CS+-CS-) in all a priori ROIs in the AA versus AC group was examined using an independent samples t-test. To address Aim 2, mean beta coefficients for each condition (CS+, CS-) in all ROIs in the extinction task were entered into a Group (AA, AC) by Condition (CS+, CS-) repeated-measures ANOVA. As in the AA/AC task, differentiation between the CS+ and CS- (quantified as CS+-CS-) in all ROIs in the extinction task was also compared in the AA versus AC group using an independent samples t-test. To establish group equivalency during acquisition, mean beta coefficients for the difference between CS+ and CS- (quantified as CS+-CS-) in the BNST, CeA, LA and BLA ROIs were compared in the AA versus AC group. To correct for multiple comparisons, Benjamini-Hochberg's adjustment was used where necessary (Benjamini & Hochberg, 1995).

Associations with individual differences in anxiety. To examine whether individual differences in trait anxiety is associated with differences in neural activation during extinction (Aim 3), mean beta coefficients for CS+ versus CS- (quantified as CS+-CS-) for each a priori ROI were correlated with STAI-T total scores. Correction for multiple comparisons was

completed using Benjamini-Hochberg's adjustment where necessary (Benjamini & Hochberg, 1995).

Results

Participant characteristics

Participant characteristics are provided in *Table 1*. There was a significant difference in age between women and men, $t(51) = 2.468$, $p < 0.05$, such that the men were older ($M = 23.79$, $SD = 4.24$) than women ($M = 21.32$, $SD = 3.00$). There were no significant differences in self-reported trait anxiety between men and women. There were no significant differences in gender, age, or trait anxiety of participants in the AA versus AC group.

	Mean (SD) or %
Gender	
Women	64.2%
Men	35.8%
Age	22.21 (3.66)
STAI Trait Anxiety	37.28 (8.06)

Table 1. Sample Characteristics. STAI, Spielberger State-Trait Anxiety Inventory.

Fear acquisition task fMRI

Voxelwise analysis. Results of a paired sample t-test for all participants (collapsed across AA and AC groups) indicated significantly greater bilateral insula (left: 38 voxels, 31.5, -19.5, -1.5, $t = 4.125$, $df = 48$, $p < 0.001$; right: 83 voxels, -35.5, 120.5, -2.5, $t = 4.498$, $df = 48$, $p < 0.001$) and right-hemisphere caudate (27 voxels, -14.5, 5.5, 22.5, $t = 4.580$, $df = 48$, $p < 0.001$) activation to the CS+ versus CS- during early-phase acquisition. This increased activation to the CS+ versus CS- was preserved in late-phase acquisition in the left-hemisphere insula (left: 21 voxels, 38.5, -15.5, -4.5, $t = 3.927$, $df = 48$, $p < 0.001$) and right-hemisphere caudate (21 voxels, -21.5, -21.5, 16.5, $t = 4.21$, $df = 48$, $p < 0.001$). Significantly decreased left-hemisphere

putamen (48 voxels, 25.5, 8.5, 3.5, t = -4.025, df = 48, $p < 0.001$) and right-hemisphere hippocampus (40 voxels, -17.5, 9.5, -14.5, t = -4.040, df = 48, $p < 0.001$) activation to the CS+ compared to the CS- during early-phase acquisition was also observed, with decreased activation to the CS+ in the bilateral hippocampus observed in late-phase acquisition (left: 23 voxels, 24.5, 28.5, -6.5, t = -4.261, df = 48, $p < 0.001$; right: 56 voxels, -27.5, 19.5, -16.5, t = -4.176, df = 48, $p < 0.001$). Results of an independent samples t-test indicated no significant differences in differentiation between the CS+ and CS- in the AA group compared to the AC group in early- or late-phase acquisition.

ROI analysis. Results of a repeated measures Group (AA, AC) by Condition (CS+, CS-) ANOVA revealed a significant Group by Condition interaction in the CeA during late-phase acquisition, such that participants in the AA group showed greater CeA activation to the CS+ compared to the AC group, $F(1) = 7.33$, $p < 0.01$. There were no significant differences in activation to the CS+ versus CS- in the remaining a priori ROIs in early- or late-phase acquisition. Independent samples t-tests were conducted in all a priori ROIs to compare differentiation between the CS+ and CS- in the AA versus AC group. There was a significant group difference in differentiation between the CS+ and CS- in the CeA during late-phase acquisition, such that participants in the AA group showed less differentiation between the CS+ and CS- than participants in the AC group $t(46) = 2.708$, $p < 0.01$. These results are shown in *Figure 3.* No other significant differences in differentiation between the CS+ and CS- in the a priori ROIs were found in early- or late-phase acquisition.

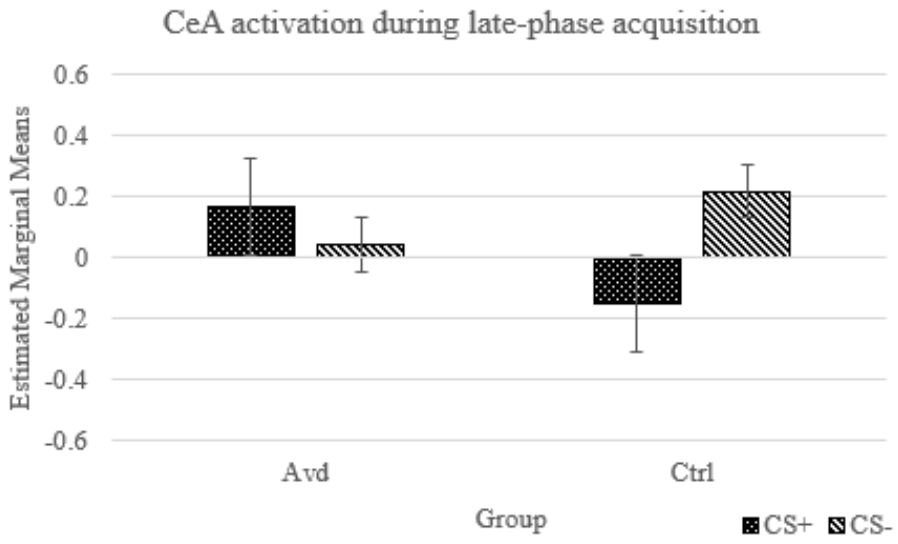

Figure 3. Central amygdala of the nucleus activation to the CS+ and CS- during late-phase acquisition in AA versus AC groups.

AA task versus AC task fMRI

Voxelwise analysis. Results of a series of independent samples t-tests indicated significantly less activation in the right-hemisphere putamen to the CS+ (76 voxels, -17.5, -16.5, -4.5, t = -4.972, df = 48, $p < 0.001$) during early-phase, but not late-phase, AA versus AC. This result is shown in *Figure 4.* There were no significant differences in differentiation between the CS+ and CS- in the AA group compared to the AC group in early- or late-phase AA/AC.

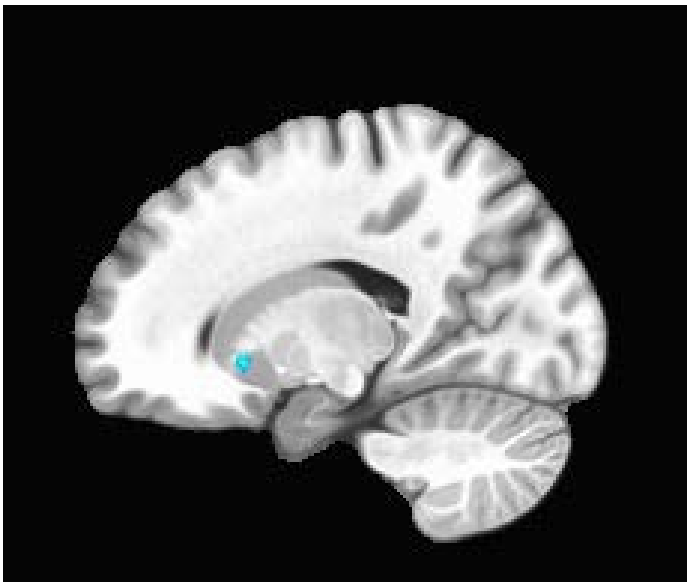

Figure 4. BOLD activation to the CS+ in the right-hemisphere putamen during early-phase AA versus AC (AA<AC).

ROI analysis. Results of a repeated measures Group (AA, AC) by Condition (CS+, CS-) ANOVA indicated a significant main effect of condition in the left-hemisphere NAcc ($F = 4.219$, $df = 1$, $p = 0.045$), such that participants in both groups showed decreased activation to the CS+ compared to the CS- during late-phase AA/AC. This finding is reported in *Figure 5.* Results of an independent samples t-test indicated no significant differences in differentiation between the CS+ and CS- in the a priori ROIs in the AA group compared to the AC group in early- or late-phase AA/AC.

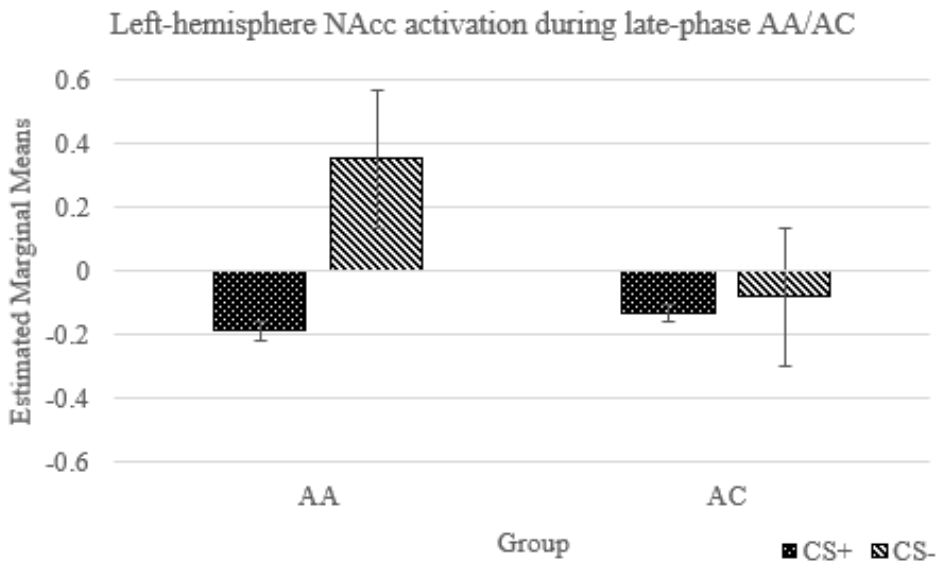

Figure 5. Left-hemisphere nucleus accumbens activation to the CS+ and CS- during late-phase AA and AC.

Extinction task fMRI in AA group versus AC group

Voxelwise analysis. Results of an independent samples t-test revealed significantly greater activation to the CS+ in the left-hemisphere putamen during late-phase extinction in the AA compared to AC group (77 voxels, 20.5, -7.5, -10.5, $t = 4.454$, $df = 49$, $p = 0.001$). This finding is shown in *Figure 6.* Further, results of another independent samples t-test indicated greater differentiation between the CS+ and CS- during early- and late-phase extinction in the left-hemisphere putamen in the AA versus AC groups (early-phase: 22 voxels, 19.5, -8.5, 5.5, $t = 4.315$, $df = 49$, $p = 0.001$; late-phase: 80 voxels, 28.5, 14.5, 11.5, $t = 5.012$, $df = 49$, $p = 0.001$). These findings are reported in *Figure 7*.

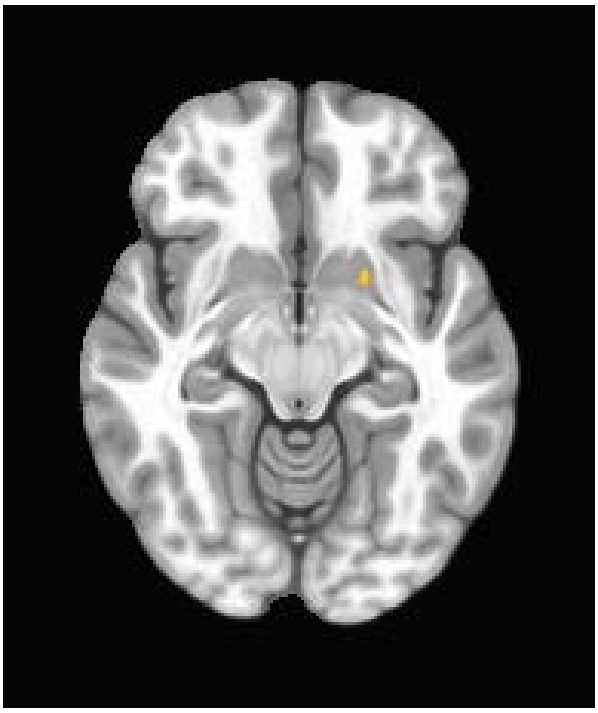

Figure 6. BOLD activation to the CS+ in the left-hemisphere putamen during late-phase extinction (AA>AC).

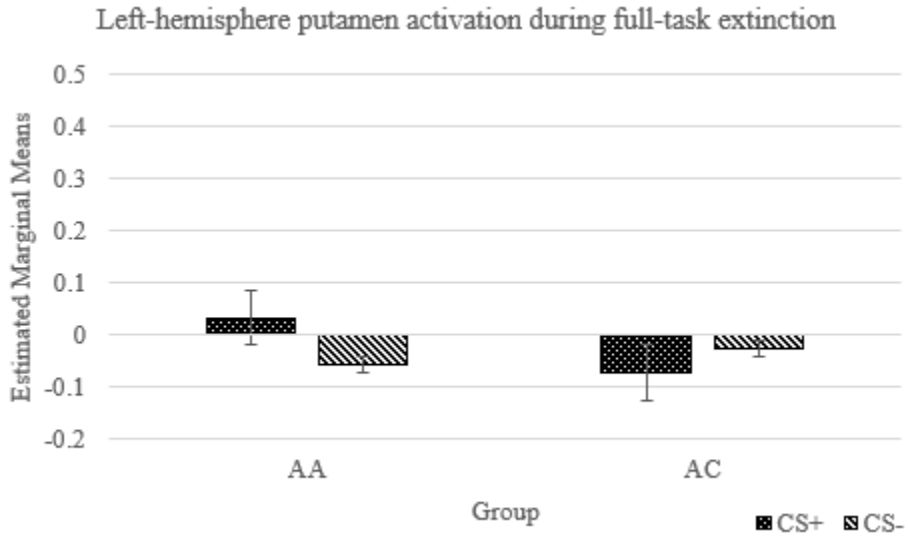

Figure 7. Left-hemisphere putamen activation to the CS+ and CS- in the AA and AC groups during extinction (collapsed across early- and late-phase).

ROI findings. Results of a repeated measures Group (AA, AC) by Condition (CS+, CS-) ANOVA indicated a significant interaction in the left-hemisphere putamen ($F = 4.286$, $df = 1$, $p = 0.044$) and right-hemisphere habenula ($F = 4.212$, $df = 1$, $p = 0.045$) during late-phase extinction. Results of an independent samples t-test revealed, also in late-phase extinction,

significantly greater differentiation between the CS+ and CS- in the left-hemisphere putamen (t = 2.070, df = 49, p = 0.044) and right-hemisphere habenula (t = 2.052, df = 49, p = 0.045) in the AA group compared to the AC group. The right-hemisphere habenula finding is reported in *Figure 8*.

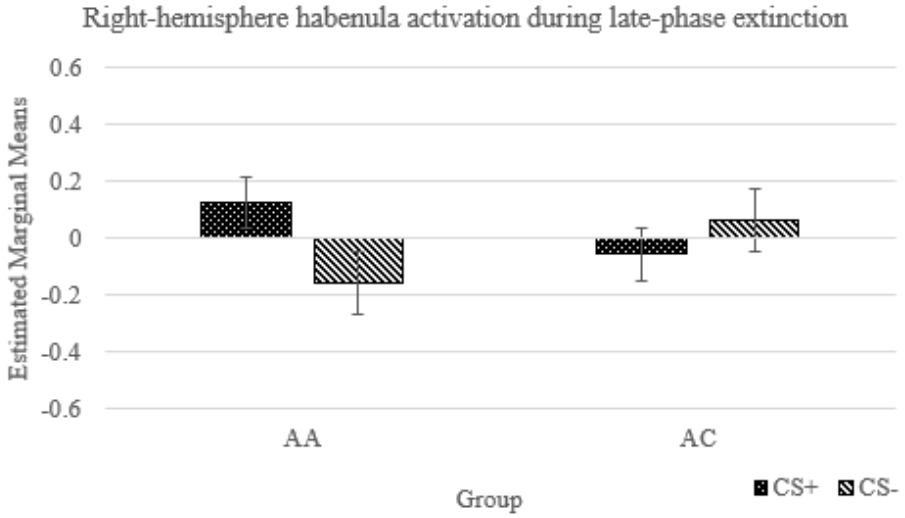

Figure 8. Right-hemisphere habenula activation to the CS+ and CS- in late-phase extinction.

Association between trait anxiety and neural activation during fear extinction

Results of the Pearson correlations indicated no significant association between participants' self-reported trait anxiety and differentiation between the CS+ and CS- in a priori ROIs during the early- or late- extinction phase for the AA and AC groups. Results are presented in *Table 2*.

	Putamen (R)	Putamen (L)	Habenula (R)	Habenula (L)	Caudate (R)	Caudate (L)	Accumbens (R)	Accumbens (L)	BNST	BLA	CeA	LA
Early-Phase Extinction (AC)	-.298	-.249	-.251	-.240	-.324	-.291	-.374	-.166	-.158	.213	-.109	.022
Early-Phase Extinction (AA)	.182	.162	.216	.224	.188	.121	-.184	-.104	-.019	.038	-.143	.009
Late-Phase Extinction (AC)	.163	.189	-.014	.048	.227	.277	-.035	-.035	.063	.083	.055	.087
Late-Phase Extinction (AA)	.281	.235	-.281	-.199	.222	.311	-.188	-.153	-.212	.180	.191	.149

Table 2. Pearson correlations between STAI-T total scores and mean beta coefficients for CS+-CS- in a priori ROIs. All *ps* = ns.

Discussion

The goal of this study was to examine neural activation to threat cues during avoidance and differences in extinction learning between groups exposed to avoidance versus non-avoidance. Avoidance is a prominent and maladaptive feature of anxiety disorders, and there is empirical support for the theory that avoidance responding interferes with extinction by preventing exposure to stimuli/events perceived as aversive. However, there is a lack of evidence from neuroimaging studies to support this theory. Importantly, this study sought to clarify the discrepancy between previous findings suggesting that avoidance impairs extinction learning and findings of a neuroimaging study by Boeke et al. (2017) suggesting that avoidance enhances fear reduction relative to extinction.

During the avoidance/control portion of the task, in early-phase AA versus AC, attenuated activation to the CS+ and less differentiation between the CS+ and CS- in the right hemisphere putamen was apparent. The putamen, part of the dorsal striatum, is involved in reward-associated learning and acquisition of goal-directed behavior (Brovelli, Nazarian, Meunier & Boussaoud, 2011). Although increased activity in the putamen was expected in the AA group, it is possible that due to the nature of the AA task (i.e., instructed avoidance) participants in the AC group (i.e., extinction) were engaging in more motivationally-salient learning than participants in the AA group during the first half of the task (e.g., learning that the CS+ is now safe). Results from the ROI analyses also indicated a main effect of Condition for the NAcc, such that participants in both the AA and AC groups showed attenuated activation in the NAcc in response to the CS+ compared to the CS- during late-phase AA/AC. However, this effect appeared to be driven by increased NAcc activation to the CS- in the AA group during late-phase avoidance. Part of the ventral striatum, the NAcc is implicated in reward-related

processing. Some studies suggest that activation in the NAcc is potentiated in response to reward prediction errors (e.g., when an expected reward is not obtained/delivered) (Magno et al., 2006; Bray & O'Doherty, 2007). In line with this, decreased activation in the nucleus accumbens during CS+ trials may suggest that viewing the CS+ is processed as more rewarding than viewing the CS- in late-phase AA/AC. Considering that participants in the AC group are undergoing extinction (i.e., the US is no longer paired with presentations of the CS+), it could be expected that the motivational salience of the CS+ is altered during late-phase AC. It is also reasonable to conclude that by late-phase AA, control over delivery of the US through active avoidance during presentation of the CS+ may be processed as more rewarding than the same behavioral response when it has no motivational significance (i.e., during presentation of the CS-). If cancellation of the US through active avoidance is processed as rewarding, this provides some support for the hypothesis that active avoidance is maintained through negative reinforcement. However, it should be noted that a similar effect was observed in the AC group, which suggests that extinction of the US-CS+ association may also be processed as rewarding. Further, the literature of the role of the NAcc in reward processing is mixed, with some studies emphasizing the role of the NAcc in reward prediction and anticipation rather than processing reward outcomes (Knutson & Cooper, 2005). Therefore, alternative explanations for these findings are possible and they should be interpreted with caution.

The second aim of this project was to examine the effect of active avoidance on subsequent extinction of conditioned fear. In both early- and late-phase extinction, participants in the AA group showed greater differentiation between the CS+ and CS- in the left-hemisphere putamen compared to the AC group. In late-phase extinction only, participants in the AA group showed significantly potentiated left-hemisphere putamen activation to the CS+ relative to the

AC group. These findings are particularly interesting considering that the opposite effect was found during the avoidance (AA) or control (AC) tasks themselves, such that participants in the AA group showed attenuated left-hemisphere putamen activation in response to the CS+ and less differentiation between the CS+ and CS- compared to the AC group. A possible explanation for this effect is that while participants in the AC group have already undergone extinction (i.e., received presentations of the CS+ without the US), participants in the AA group are now undergoing extinction for the first time. As seen in the AC group in the previous task, participants in the AA group may show greater potentiation of the left-hemisphere putamen because of the motivationally salient learning that is taking place, specifically extinction of the US-CS+ contingency and avoidance response. Indeed, there is evidence to suggest that, in addition to its role in reward prediction, the putamen plays a role in predictive fear learning by controlling allocation of attention to predictors of danger (McNally & Westbrook, 2006). Further, some studies have implicated the putamen in explicit inhibitory control and implicit emotion regulation (Dibbets et al., 2010; Jarcho et al., 2013). Thus, it is possible that increased activation in the putamen is indicative of updating of the US-CS contingency and suppression of the conditioned response.

In addition, participants in the AA group showed greater differentiation between the CS+ and CS- in the right-hemisphere habenula relative to the AC group during late-phase extinction. Although the role of the habenula in motivation and decision-making is not well-studied in humans, studies in non-human vertebrates suggest that it may play a critical role in modulation of motivated behavior and reward prediction, as well as regulation of fear expression (Namboodiri et al., 2016). Thus, increased differentiation between the CS+ and CS- in the right-hemisphere habenula may be indicative of successful suppression of the conditioned response

and avoidance behavior. As mentioned above, evidence of differentiation between the CS+ and CS- in the putamen and habenula in the AA group and not the AC group can be accounted for by the fact that the active control condition is essentially an extinction task (the CS+ is no longer paired with the US, and instructed motor responses are inconsequential). Thus, during the subsequent extinction task, participants in the AC group have already underwent extinction. Alternatively, participants in the AA group did not undergo extinction in the active avoidance task (instructed motor responses are used to avoid the US, so the US-CS+ contingency remains) and are undergoing extinction of the US-CS+ contingency for the first time in the extinction task.

The findings of increased activation to the CS+ and discrimination between the CS+ and CS- in regions involved in motivational learning and emotion regulation in the AA group during extinction conflict with the a priori hypothesis that active avoidance would impair subsequent extinction learning. Notably, previous studies whose findings suggest a rebound of conditioned responding and "protection from extinction" following active avoidance have relied exclusively on peripheral physiological responses (e.g., skin conductance) and threat expectancy ratings (Lovibond et al., 2000; Lovibond et al., 2009; Vervliet & Indekeu, 2015). While peripheral physiological responses are often used as a proxy for neural activity, it is generally accepted that some are indices of physiological arousal rather than perceived stimulus valence, and thus also index appetitive responses (e.g., skin conductance; Bradley, Codispoti, Cuthbert & Lang, 2001). Additionally, neuroimaging studies on the generation and representation of the skin conductance response (SCR) at the neural level suggest that BOLD activation in regions implicated in motivated behavior and fear processing does not necessarily covary with changes in skin conductance (Critchley, Elliott, Mathias & Dolan, 2000).

Notably, Boeke et al. (2017) did not replicate previous SCR findings of reinstatement of fear following avoidance. The authors reported no significant difference in SCR to the CS+-CS- in participants who engaged in active avoidance versus yoked extinction during these tasks and in a subsequent extinction retrieval task. Interestingly, they reported a significant between-group difference in the bilateral putamen during avoidance/yoked extinction tasks, such that participants in the active avoidance group showed greater bilateral putamen activation to the CS+ compared to participants in the yoked extinction group. In the current study, results indicated that participants in the AC group showed greater activation in the putamen to the CS+ compared to the AA group during the avoidance/control task. A possible explanation for this discrepancy may be that the current study utilized an instructed avoidance paradigm, while participants in the avoidance group in Boeke et al. (2017) underwent associative avoidance learning. Considering the role of the putamen in reward-associated learning and acquisition of goal-directed behavior, it may be more robustly recruited when avoidance behavior is learned rather than instructed. Similar to the current study, Boeke et al. (2017) reported no between-group differences in BOLD activity during an extinction retrieval task suggestive of "protection from extinction". It is possible that, in line with the findings of Boeke et al. (2017), active avoidance results in diminishment of the conditioned response in subsequent extinction rather than a resurgence of threat responding or insufficient extinction.

Although it was expected that a clinically-relevant individual difference characteristic, trait anxiety, would relate to group differences in the extinction task, no significant correlations between trait anxiety and discrimination between the CS+ and CS- during extinction was found. This may be partially due to sample characteristics, such that the average total STAI score (37) in our sample is commonly classified as indicative of "no or low anxiety", with one standard

deviation above the mean falling into the “moderate anxiety” range (38-44; Kayikcioglu, Bilgin, Seymenoglu, & Deveci, 2017). Due to the nature of the study (e.g., receiving shocks, laying in an MRI machine), it is likely that participants high in trait anxiety were less likely to volunteer to participate or complete all study procedures.

There were notable limitations in the current study that should be considered. First, there were between-group differences observed in CeA activation in the acquisition phase. This suggests that there was not complete group equivalency at baseline and is a potential confound of the results. Second, active avoidance in the experimental task (AA) was instructed such that participants did not undergo associative avoidance learning. Concurrently, the comparison group underwent extinction learning in the AC task. This may represent a confound, such that participants in the AA group did not have a chance to engage in new learning like those in the AC group. Related to this point, while the active avoidance paradigm used in this study is commonly utilized in animal and human studies on threat avoidance, future work would likely benefit from utilizing experimental paradigms that simulate real-world behavioral avoidance by including response choice and manipulation of threat certainty. Finally, the current study utilized a convenience sample comprised of healthy college students. To investigate neural systems underlying maladaptive avoidance and its potential mechanistic role in maintaining pathological anxiety, it would likely be beneficial to sample from populations with clinical levels of internalizing symptoms.

Overall, the results of the current study suggest that active avoidance is associated with increased activity in some regions involved in reward prediction, which may represent a mechanism that maintains avoidance responses. However, this did not differentiate active avoidance from an active control condition. Additionally, neural activation during the extinction

task appeared to support extinction learning and fear suppression in participants who previously engaged in active avoidance, suggesting that behavioral avoidance did not result in “protection from extinction”. Future work attempting to disentangle the neural underpinnings of active avoidance would likely benefit from using a multimethod approach, including neuroimaging, peripheral psychophysiology (e.g., skin conductance) and self-reported threat expectancy and/or anxiety ratings. Additionally, in order to gain a mechanistic understanding of the development and maintenance of maladaptive avoidance in psychopathology, there is a need for tasks that index avoidance in a more ecologically valid manner (e.g., involving choice and threat uncertainty).

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)* American Psychiatric Pub.

Avery, S. N., Clauss, J. A., & Blackford, J. U. (2016). The human BNST: Functional role in anxiety and addiction. *Neuropsychopharmacology, 41*(1), 126.

Avery, S. N., Clauss, J. A., Winder, D. G., Woodward, N., Heckers, S., & Blackford, J. U. (2014). BNST neurocircuitry in humans. *NeuroImage, 91*, 311-323.

Barnes, L. L., Harp, D., & Jung, W. S. (2002). Reliability generalization of scores on the spielberger state-trait anxiety inventory. *Educational and Psychological Measurement, 62*(4), 603-618.

Barrera, T. L., & Norton, P. J. (2009). Quality of life impairment in generalized anxiety disorder, social phobia, and panic disorder. *Journal of Anxiety Disorders, 23*(8), 1086-1090.

Bechara, A., Tranel, D., Damasio, H., Adolphs, R., Rockland, C., & Damasio, A. R. (1995). Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science, 269*(5227), 1115-1118.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological), 57*(1), 289-300.

Blakey, S. M., & Abramowitz, J. S. (2016a). The effects of safety behaviors during exposure therapy for anxiety: Critical analysis from an inhibitory learning perspective. *Clinical Psychology Review, 49*, 1-15.

Blakey, S. M., & Abramowitz, J. S. (2016b). The effects of safety behaviors during exposure therapy for anxiety: Critical analysis from an inhibitory learning perspective. *Clinical Psychology Review, 49*, 1-15.

Boeke, E. A., Moscarello, J. M., LeDoux, J. E., Phelps, E. A., & Hartley, C. A. (2017). Active avoidance: Neural mechanisms and attenuation of pavlovian conditioned responding. *Journal of Neuroscience, 37*(18), 4808-4818.

Bolstad, I., Andreassen, O. A., Reckless, G. E., Sigvartsen, N. P., Server, A., & Jensen, J. (2013). Aversive event anticipation affects connectivity between the ventral striatum and the orbitofrontal cortex in an fMRI avoidance task. *PLoS One, 8*(6), e68494.

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: defensive and appetitive reactions in picture processing. *Emotion, 1*(3), 276.

Bray, S., & O'Doherty, J. (2007). Neural coding of reward-prediction error signals during classical conditioning with attractive faces. *Journal of neurophysiology, 97*(4), 3036-3045.

Brooks, S. J., & Stein, D. J. (2015). A systematic review of the neural bases of psychotherapy for anxiety and related disorders. *Dialogues in Clinical Neuroscience, 17*(3), 261.

Bryant, R. A., Kemp, A. H., Felmingham, K. L., Liddell, B., Olivieri, G., Peduto, A., . . . Williams, L. M. (2008). Enhanced amygdala and medial prefrontal activation during nonconscious processing of fear in posttraumatic stress disorder: An fMRI study. *Human Brain Mapping, 29*(5), 517-523.

Buhr, K., & Dugas, M. J. (2002). The intolerance of uncertainty scale: Psychometric properties of the english version. *Behaviour Research and Therapy, 40*(8), 931-945.

Bystritsky, A., Khalsa, S. S., Cameron, M. E., & Schiffman, J. (2013). Current diagnosis and treatment of anxiety disorders. *Pharmacy and Therapeutics, 38*(1), 30.

Carleton, R. N., Mulvogue, M. K., Thibodeau, M. A., McCabe, R. E., Antony, M. M., & Asmundson, G. J. (2012). Increasingly certain about uncertainty: Intolerance of uncertainty across anxiety and depression. *Journal of Anxiety Disorders, 26*(3), 468-479.

Carlson, J. M., Cha, J., & Mujica-Parodi, L. R. (2013). Functional and structural amygdala–anterior cingulate connectivity correlates with attentional bias to masked fearful faces. *Cortex, 49*(9), 2595-2600.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology, 67*(2), 319.

Corcoran, K. A., Desmond, T. J., Frey, K. A., & Maren, S. (2005). Hippocampal inactivation disrupts the acquisition and contextual encoding of fear extinction. *Journal of Neuroscience, 25*(39), 8978-8987.

Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, 29*(3), 162-173.

Critchley, H. D., Elliott, R., Mathias, C. J., & Dolan, R. J. (2000). Neural activity relating to generation and representation of galvanic skin conductance responses: a functional magnetic resonance imaging study. *Journal of Neuroscience, 20*(8), 3033-3040.

Davis, M. (1998). Are different parts of the extended amygdala involved in fear versus anxiety? *Biological Psychiatry, 44*(12), 1239-1247.

Davis, M., Walker, D. L., Miles, L., & Grillon, C. (2010). Phasic vs sustained fear in rats and humans: Role of the extended amygdala in fear vs anxiety. *Neuropsychopharmacology, 35*(1), 105.

Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D. C., & Fiez, J. A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *Journal of Neurophysiology, 84*(6), 3072-3077.

Dibbets, P., Evers, E. A., Hurks, P. P., Bakker, K., & Jolles, J. (2010). Differential brain activation patterns in adult attention-deficit hyperactivity disorder (ADHD) associated with task switching. *Neuropsychology, 24*(4), 413.

Dunsmoor, J. E., Bandettini, P. A., & Knight, D. C. (2007). Impact of continuous versus intermittent CS-UCS pairing on human brain activation during pavlovian fear conditioning.*Behavioral Neuroscience, 121*(4), 635.

Eldar, E., Hauser, T. U., Dayan, P., & Dolan, R. J. (2016). Striatal structure and function predict individual biases in learning to avoid pain. *Proceedings of the National Academy of Sciences, 113*(17), 4812-4817.

Etkin, A., Egner, T., & Kalisch, R. (2011). Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends in Cognitive Sciences, 15*(2), 85-93.

Felmingham, K., Kemp, A., Williams, L., Das, P., Hughes, G., Peduto, A., & Bryant, R. (2007). Changes in anterior cingulate and amygdala after cognitive behavior therapy of posttraumatic stress disorder. *Psychological Science, 18*(2), 127-129.

Freeston, M. H., Rhéaume, J., Letarte, H., Dugas, M. J., & Ladouceur, R. (1994). Why do people worry? *Personality and Individual Differences, 17*(6), 791-802.

Grady, A. K., Bowen, K. H., Hyde, A. T., Totsch, S. K., & Knight, D. C. (2016). Effect of continuous and partial reinforcement on the acquisition and extinction of human conditioned fear.*Behavioral Neuroscience, 130*(1), 36.

Hariri, A. R., Tessitore, A., Mattay, V. S., Fera, F., & Weinberger, D. R. (2002). The amygdala response to emotional stimuli: A comparison of faces and scenes. *NeuroImage, 17*(1), 317-323.

Ilango, A., Kesner, A. J., Keller, K. L., Stuber, G. D., Bonci, A., & Ikemoto, S. (2014a). Similar roles of substantia nigra and ventral tegmental dopamine neurons in reward and aversion. *Journal of Neuroscience, 34*(3), 817-822.

Ilango, A., Shumake, J., Wetzel, W., & Ohl, F. W. (2014b). Contribution of emotional and motivational neurocircuitry to cue-signaled active avoidance learning. *Frontiers in Behavioral Neuroscience, 8*, 372.

Jarcho, J. M., Fox, N. A., Pine, D. S., Etkin, A., Leibenluft, E., Shechner, T., & Ernst, M. (2013). The neural correlates of emotion-based cognitive control in adults with early childhood behavioral inhibition. *Biological psychology, 92*(2), 306-314.

Jensen, J., McIntosh, A. R., Crawley, A. P., Mikulis, D. J., Remington, G., & Kapur, S. (2003). Direct activation of the ventral striatum in anticipation of aversive stimuli. *Neuron, 40*(6), 1251-1257.

Johnstone, T., Somerville, L. H., Alexander, A. L., Oakes, T. R., Davidson, R. J., Kalin, N. H., & Whalen, P. J. (2005). Stability of amygdala BOLD response to fearful faces over multiple scan sessions. *NeuroImage, 25*(4), 1112-1123.

Jorm, A. F., Christensen, H., Henderson, A. S., Jacomb, P. A., Korten, A. E., & Rodgers, B. (1998). Using the BIS/BAS scales to measure behavioural inhibition and behavioural activation: Factor structure, validity and norms in a large community sample. *Personality and Individual Differences, 26*(1), 49-58.

Kayikcioglu, O., Bilgin, S., Seymenoglu, G., & Deveci, A. (2017). State and Trait Anxiety Scores of Patients Receiving Intravitreal Injections. *Biomedicine Hub, 2*(2), 1-5.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Archives of General Psychiatry, 62*(6), 593-602.

Kessler, R. C., Ruscio, A. M., Shear, K., & Wittchen, H. (2009). Epidemiology of anxiety disorders. *Behavioral neurobiology of anxiety and its treatment* (pp. 21-35) Springer.

Knutson, B., & Cooper, J. C. (2005). Functional magnetic resonance imaging of reward prediction. *Current opinion in neurology, 18*(4), 411-417.

LaBar, K. S., Gatenby, J. C., Gore, J. C., LeDoux, J. E., & Phelps, E. A. (1998). Human amygdala activation during conditioned fear acquisition and extinction: A mixed-trial fMRI study.*Neuron, 20*(5), 937-945.

Lebow, M. A., & Chen, A. (2016). Overshadowed by the amygdala: The bed nucleus of the stria terminalis emerges as key to psychiatric disorders. *Molecular Psychiatry, 21*(4), 450.

Levita, L., Hoskin, R., & Champi, S. (2012). Avoidance of harm and anxiety: A role for the nucleus accumbens. *NeuroImage, 62*(1), 189-198.

Lissek, S., Powers, A. S., McClure, E. B., Phelps, E. A., Woldehawariat, G., Grillon, C., & Pine, D. S. (2005). Classical fear conditioning in the anxiety disorders: A meta-analysis. *Behaviour Research and Therapy, 43*(11), 1391-1424.

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*(1), 9-16.

Lovibond, P. F., Davis, N. R., & O'Flaherty, A. S. (2000). Protection from extinction in human fear conditioning. *Behaviour Research and Therapy, 38*(10), 967-983.

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*(8), 716-720.

Makinson, R. A., & Young, J. S. (2012). Cognitive behavioral therapy and the treatment of posttraumatic stress disorder: Where counseling and neuroscience meet. *Journal of Counseling & Development, 90*(2), 131-140.

Magno, E., Foxe, J. J., Molholm, S., Robertson, I. H., & Garavan, H. (2006). The anterior cingulate and error avoidance. *Journal of Neuroscience, 26*(18), 4769-4773.

Marschner, A., Kalisch, R., Vervliet, B., Vansteenwegen, D., & Büchel, C. (2008). Dissociable roles for the hippocampus and the amygdala in human cued versus context fear conditioning.*Journal of Neuroscience, 28*(36), 9030-9036.

McNally, R. J. (2007). Mechanisms of exposure therapy: How neuroscience can improve psychological treatments for anxiety disorders. *Clinical Psychology Review, 27*(6), 750-759.

McNally, G. P., & Westbrook, R. F. (2006). Predicting danger: the nature, consequences, and neural mechanisms of predictive fear learning. *Learning & Memory, 13*(3), 245-253.

Myers, K. M., & Davis, M. (2007). Mechanisms of fear extinction. *Molecular Psychiatry, 12*(2), 120.

Namboodiri, V. M. K., Rodriguez-Romaguera, J., & Stuber, G. D. (2016). The habenula. *Current Biology, 26*(19), R873-R877.

Norton, P. J., Temple, S. R., & Pettit, J. W. (2008). Suicidal ideation and anxiety disorders: Elevated risk or artifact of comorbid depression? *Journal of Behavior Therapy and Experimental Psychiatry, 39*(4), 515-525.

Öhman, A. (2002). Automaticity and the amygdala: Nonconscious responses to emotional faces. *Current Directions in Psychological Science, 11*(2), 62-66.

Öhman, A., Esteves, F., & Soares, J. J. (1995). Preparedness and preattentive associative learning: Electrodermal conditioning to masked stimuli. *Journal of Psychophysiology,*

Pessoa, L., Japee, S., & Ungerleider, L. G. (2005). Visual awareness and the detection of fearful faces. *Emotion, 5*(2), 243.

Pohlack, S. T., Nees, F., Ruttorf, M., Schad, L. R., & Flor, H. (2012). Activation of the ventral striatum during aversive contextual conditioning in humans. *Biological Psychology, 91*(1), 74-80.

Quirk, G. J., Garcia, R., & González-Lima, F. (2006). Prefrontal mechanisms in extinction of conditioned fear. *Biological Psychiatry, 60*(4), 337-343.

Quirk, G. J., & Mueller, D. (2008). Neural mechanisms of extinction learning and retrieval. *Neuropsychopharmacology, 33*(1), 56.

Rauch, S. L., Shin, L. M., & Phelps, E. A. (2006). Neurocircuitry models of posttraumatic stress disorder and extinction: Human neuroimaging research—past, present, and future.*Biological Psychiatry, 60*(4), 376-382.

Sareen, J., Cox, B. J., Afifi, T. O., de Graaf, R., Asmundson, G. J., Ten Have, M., & Stein, M. B. (2005). Anxiety disorders and risk for suicidal ideation and suicide attempts: A population-based longitudinal study of adults. *Archives of General Psychiatry, 62*(11), 1249-1257.

Schlund, M. W., Brewer, A. T., Magee, S. K., Richman, D. M., Solomon, S., Ludlum, M., & Dymond, S. (2016). The tipping point: Value differences and parallel dorsal–ventral frontal circuits gating human approach–avoidance behavior. *NeuroImage, 136*, 94-105.

Schlund, M. W., & Cataldo, M. F. (2010). Amygdala involvement in human avoidance, escape and approach behavior. *NeuroImage, 53*(2), 769-776.

Schlund, M. W., Magee, S., & Hudgins, C. D. (2011). Human avoidance and approach learning: Evidence for overlapping neural systems and experiential avoidance modulation of avoidance neurocircuitry. *Behavioural Brain Research, 225*(2), 437-448.

Schlund, M. W., Siegle, G. J., Ladouceur, C. D., Silk, J. S., Cataldo, M. F., Forbes, E. E., . . . Ryan, N. D. (2010). Nothing to fear? neural systems supporting avoidance behavior in healthy youths. *NeuroImage, 52*(2), 710-719.

Seymour, B., Daw, N., Dayan, P., Singer, T., & Dolan, R. (2007). Differential encoding of losses and gains in the human striatum. *Journal of Neuroscience, 27*(18), 4826-4831.

Shackman, A. J., & Fox, A. S. (2016). Contributions of the central extended amygdala to fear and anxietycontributions of the central extended amygdala to fear and anxiety. *Journal of Neuroscience, 36*(31), 8050-8063.

Sheynin, J., Beck, K. D., Pang, K. C., Servatius, R. J., Shikari, S., Ostovich, J., & Myers, C. E. (2014). Behaviourally inhibited temperament and female sex, two vulnerability factors for anxiety disorders, facilitate conditioned avoidance (also) in humans. *Behavioural Processes, 103*, 228-235.

Shumake, J., Ilango, A., Scheich, H., Wetzel, W., & Ohl, F. W. (2010). Differential neuromodulation of acquisition and retrieval of avoidance learning by the lateral habenula and ventral tegmental area. *Journal of Neuroscience, 30*(17), 5876-5883.

Sotres-Bayon, F., Cain, C. K., & LeDoux, J. E. (2006). Brain mechanisms of fear extinction: Historical perspectives on the contribution of prefrontal cortex. *Biological Psychiatry, 60*(4), 329-336.

Spielberger, C. D., Sydeman, S. J., Owen, A. E., & Marsh, B. J. (1999). Measuring anxiety and anger with the state-trait anxiety inventory (STAI) and the state-trait anger expression inventory (STAXI).

Spielberger, C. D., & Gorsuch, R. L. (1983). *State-trait anxiety inventory for adults: Sampler set: Manual, test, scoring key.* Mind Garden.

Stein, M. B., Roy-Byrne, P. P., Craske, M. G., Bystritsky, A., Sullivan, G., Pyne, J. M., . . . Sherbourne, C. D. (2005). Functional impact and health utility of anxiety disorders in primary care outpatients. *Medical Care*, 1164-1170.

Sullivan, G. M., Apergis, J., Bush, D., Johnson, L. R., Hou, M., & Ledoux, J. E. (2004). Lesions in the bed nucleus of the stria terminalis disrupt corticosterone and freezing responses elicited by a contextual but not by a specific cue-conditioned fear stimulus. *Neuroscience, 128*(1), 7-14.

Tyszka, J. M., & Pauli, W. M. (2016). In vivo delineation of subdivisions of the human amygdaloid complex in a high-resolution group template. *Human Brain Mapping, 37*(11), 3979-3998.

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

Volders, S., Meulders, A., De Peuter, S., Vervliet, B., & Vlaeyen, J. W. (2012). Safety behavior can hamper the extinction of fear of movement-related pain: An experimental investigation in healthy participants. *Behaviour Research and Therapy, 50*(11), 735-746.

Whalen, P. J., Kagan, J., Cook, R. G., Davis, F. C., Kim, H., Polis, S., . . . Maxwell, J. S. (2004). Human amygdala responsivity to masked fearful eye whites. *Science, 306*(5704), 2061.