From Pavlov to PTSD: The extinction of conditioned fear in rodents, humans, and in anxiety disorders

Michael B. VanElzakker, M. Kathryn Dahlgren, F. Caroline Davis, Stacey Dubois, and Lisa M. Shin

aTufts University Psychology, 490 Boston Avenue, Medford, MA 02155, USA
bMassachusetts General Hospital Psychiatry, 149 Thirteenth Street, Charlestown, MA 02129, USA

Abstract

Nearly 100 years ago, Ivan Pavlov demonstrated that dogs could learn to use a neutral cue to predict a biologically relevant event: after repeated predictive pairings, Pavlov's dogs were conditioned to anticipate food at the sound of a bell, which caused them to salivate. Like sustenance, danger is biologically relevant, and neutral cues can take on great salience when they predict a threat to survival. In anxiety disorders such as posttraumatic stress disorder (PTSD), this type of conditioned fear fails to extinguish, and reminders of traumatic events can cause pathological conditioned fear responses for decades after danger has passed. In this review, we use fear conditioning and extinction studies to draw a direct line from Pavlov to PTSD and other anxiety disorders. We explain how rodent studies have informed neuroimaging studies of healthy humans and humans with PTSD. We describe several genes that have been linked to both PTSD and fear conditioning and extinction and explain how abnormalities in fear conditioning or extinction may reflect a general biomarker of anxiety disorders. Finally, we explore drug and neuromodulation treatments that may enhance therapeutic extinction in anxiety disorders.

1. INTRODUCTION

In his classical conditioning and extinction experiments, Ivan Pavlov rang a bell (the conditioned stimulus; CS), immediately before giving his dogs food (specifically meat powder, the unconditioned stimulus; US; Pavlov, 1927). On its own, the meat powder made the dogs salivate (the unconditioned response; UR). After repeating this predictive pairing several times, Pavlov's dogs began salivating to the mere sound of the bell—even when no meat powder was presented—making salivation the conditioned response (CR). The sound of the bell predicted something agreeable and biologically valuable: food. However, not all of Pavlov's USs were pleasant, and not all CRs conveyed his dogs' anticipation of something
enjoyable. In addition to learning about nourishment sources, it is important for an organism to be able to predict threats to health and safety. For example, when Pavlov repeatedly paired the sound of a metronome (CS) with subsequent application of a small amount of sour-tasting diluted acid (US) onto a dog's tongue, the dog eventually learned the association. Henceforth, upon presentation of the CS alone, the dog exhibited what Pavlov called a “defensive reflex”: it shook its head, salivated profusely, and moved its tongue as if to expel a toxic substance, even though no acid was there. A similar process was demonstrated with an 11-month-old child in Watson and Rayner's famous “Little Albert” experiments of 1920. Watson and Rayner paired Albert's touching of a white rat (CS) with a sudden fear-arousing noise (US) made by striking a steel bar behind him (Watson & Rayner, 2000). Upon subsequent presentations of the rat, Albert no longer exhibited his natural curiosity, but rather withdrew his hand. This learned response seemed to generalize to cotton balls, a Santa Claus mask, a brown bunny, and a black fur coat. The Little Albert experiment is an early precursor of what is now known as fear conditioning.

It is not known whether Little Albert subsequently experienced fear around rats and furry objects (if he survived into adulthood at all) or if he was healthy and well-adjusted (Harris, 2011). Of course, modern ethical standards would not allow such a methodology. Still, it is likely that, after the experiment was over, Little Albert encountered other rats or other furry objects in the absence of a loud noise. Eventually, he should have learned that such objects no longer predicted a frightening clang, and his fear response should have declined. This process is known as fear extinction learning. When the CS no longer predicts the US, the conditioned fear response is extinguished.

How do these processes of fear conditioning and fear extinction work? Why is it that with very severe USs, some individuals are burdened by fear and anxiety for decades? The goal of this review is to examine the underlying mechanisms and neurocircuitry of fear conditioning and extinction, as well as to explore how these processes can inform our understanding of anxiety disorders such as posttraumatic stress disorder (PTSD). We will first discuss fear conditioning and extinction in rodents, and then in healthy humans. Finally, we'll discuss fear conditioning and extinction in individuals with PTSD and other anxiety disorders, with an emphasis on how extinction learning relates to treatment.

### 2.1 FEAR CONDITIONING IN RODENTS

When rodents sense danger, one species-specific behavioral response is to freeze all movement in order to avoid detection by predators. Rodent fear conditioning and extinction studies typically use a foot shock as the US. The fear response is operationalized as the percentage of time a rodent spends engaging in freezing behavior. When a light or tone (CS) repeatedly predicts a foot shock (US) delivered through an electrified metal cage floor, rodents are conditioned to make a CS-US association. Thus, the presence of the CS subsequently triggers freezing, which becomes the CR. Furthermore, when a rodent experiences an aversive US such as shock in a certain context, subsequent re-exposure to that context can cause freezing behavior, even if the shock has not been paired with a discrete CS such as a light or tone. This type of Pavlovian fear conditioning is known as contextual fear conditioning (Rudy, Huff, & Matus-Amat, 2004). When a rodent experiences
a sudden loud noise it will startle before freezing, but if that sudden loud noise occurs during
the presentation of a danger-associated cue such as a CS or a conditioning context, the
startle reflex will be larger. This is known as a fear-potentiated startle and is another
commonly used CR (Davis, 2001). The fear-potentiated startle paradigm is advantageous for
translational research because it is not species-specific.

Researchers can link conditioned behaviors such as freezing or fear-potentiated startle to
brain activity or other fear-based physiological measures. With this simple fear conditioning
model, the neurocircuitry of fear learning and extinction has been well delineated (reviewed
in more detail in this issue XXXX and also Johnson, McGuire, Lazarus, & Palmer, 2012;
LeDoux, 2000; Maren, 2001; Rudy, 2008). Here, we will briefly review the neurocircuitry
involved in fear conditioning and extinction in rodents (see Figure 1).

The sensory experiences of the CS and US are processed in the thalamus and somatosensory
cortex, as are other sensory experiences. This information reaches the lateral amygdala via
one of two routes. A “cortical pathway” relays detailed sensory information through the
thalamus to the neocortex and hippocampus before integration and evaluation in the lateral
amygdala. However, another pathway forgoes the neocortex in the service of reaction speed.
This faster “subcortical pathway” projects a rudimentary sensory representation directly
from thalamus to the lateral and central nuclei of the amygdala. The binding together of a
conditioned CS-US association is supported by the lateral nucleus of the amygdala, which
then projects to the central amygdala, triggering autonomic and behavioral responses such as
freezing (Blair, Schafe, Bauer, Rodrigues & LeDoux, 2001; Pitkanen, 2000) and fear-
potentiated startle (Campeau & Davis, 1995). The amygdala is part of a broader
neurocircuitry that supports and modulates this process.

Conditioning and extinction of rodent freezing behavior are both modulated by medial
prefrontal cortex (mPFC) structures. The more dorsal prelimbic cortex of the rodent is
associated with the expression of conditioned fear (Burgos-Robles, Vidal-Gonzalez, &
Quirk, 2009). The prelimbic cortex acts as a fear response “accelerator” during conditioning,
while the more ventral infralimbic cortex acts as “brakes” during extinction. The infralimbic
cortex is necessary for fear conditioning responses to context (Resstel, Joca, Guimarães, &
Corrêa, 2006), probably due to its connectivity to hippocampus and amygdala (Bouton,
Westbrook, Corcoran, & Maren, 2006; Maren, Phan, & Liberzon, 2013).

The hippocampus serves the function of binding together the disparate sensory and
interoceptive elements that form a context into one conjunctive representation (Rudy &
O’Reilly, 2001). The rodent hippocampus has connections with both prelimbic and
infralimbic cortex and thus provides contextual modulation over fear responses.
Furthermore, during exploration of the environment, the hippocampus, along with associated
medial temporal cortex, serves as a functional comparator of present and past (stored)
experience (VanElzakker, Fevurly, Breindel, & Spencer, 2008). As such, it is vital to the
recognition of a context as familiar or the establishment of a context as novel. A related
function is its involvement in comparing novel cues to an existing CS, to determine if a CR
is appropriate; stimulus generalization is what led Little Albert to be wary of cues that only
moderately resembled a white rat. The hippocampus is therefore a crucial structure in
determining whether contextual cues are associated with danger or with safety (Maren, 2013; Rudy et al., 2004).

2.2 EXTINCTION IN RODENTS

At the level of behavioral observation, if a conditioned cue (CS) or context is repeatedly presented without subsequent shock (US), the rat will stop freezing in response to the CS. This process of extinction is somewhat tenuous, as its recall is fundamentally context-dependent (Bouton, 2004; Bouton, Westbrook, et al., 2006). That is, once both a CS-US (conditioning) and a CS-noUS (extinction) representation exist, the response relevant to CS-noUS is only expressed in the context in which CS-noUS was learned. Furthermore, reduction of the CR does not necessarily mean that the CS-US association has been broken. This is demonstrated by the phenomena of spontaneous recovery, reinstatement, and renewal. As Pavlov described, spontaneous recovery refers to the fact that, after the passage of time, the CS can recover the ability to elicit an extinguished CR. Reinstatement of an extinguished CR occurs when the US is presented in the absence of the CS; simple exposure to a US, even outside of the conditioning or extinction context or without being paired with any particular cue, can reinstate fear responses to a previously conditioned context or cue. More recent rodent research has also revealed the phenomenon of renewal, which occurs when conditioning and extinction occur in different contexts: a change from the extinction context either back to the conditioning context or into a third context can cause the CR to renew (Bouton, 2004). Therefore, while it may be intuitive to conclude that extinction of the CR represents a fading away of the CS-US association, these three phenomena provide evidence that fear extinction primarily represents a competing memory (Herry et al., 2010) because the CS can still recover the ability to cause a CR without ever being re-paired with the US. This demonstrates that extinction learning represents a new CS-noUS memory trace that competes with and inhibits the existing CS-US memory.

So if extinction represents new learning, how can it give rise to inhibition over a prepotent fear response? There is evidence to suggest that there is a necessary excitatory component to this process of fear response inhibition: in rodents, blocking NMDA receptors (N-methyl-D-aspartate, a receptor for the excitatory neurotransmitter glutamate) impairs extinction (Davis, 2011; Zimmermann & Maren, 2010). The rodent infralimbic cortex has direct, excitatory projections to the amygdala's intercalated neurons (Beretta, Pantazopoulos, Caldera, Pantazopoulos, & Paré, 2005), a group of GABA (gamma-aminobutyric acid, an inhibitory neurotransmitter) producing cells with inhibitory influence on amygdala output. These neurons project to the central amygdala, representing the extinction association (CS-noUS) as a countering influence against the excitatory projections from the lateral nucleus that represent the conditioned association (CS-US). Indeed, excitation of infralimbic cortex during presentation of a CS (Milad & Quirk, 2002) or in a fear context (Thompson et al., 2010) causes extinction. Thus, extinction of conditioned fear responses represents inhibition of the central amygdala output neurons that control freezing or startle behavior via their projections to the midbrain (Amano, Unal, & Paré, 2010). The infralimbic cortex is also involved in fear extinction recall (Milad & Quirk, 2012). Studies in mice have demonstrated that distinct neurons in the basal nucleus of the amygdala appear to encode the differential conditioning memory and extinction memory (Herry et al., 2008).
2.3 RECONSOLIDATION

Although the focus of this review is on extinction, we should also mention the phenomenon of memory reconsolidation. When a CS+ is presented after conditioning, the fear memory is reactivated, enters a labile state, and is open to (1) reconsolidation or (2) reconsolidation blockade using either a pharmacologic agent (e.g., Gamache, Pitman & Nader, 2012; Kindt, Soeter & Vervliet, 2009; Nader, Schafe & LeDoux, 2000; Nader & Hardt, 2009; Pitman et al., 2011) or immediate extinction within the lability window (e.g., Monfils, Cowansage, Klann & LeDoux, 2009). Repeated reconsolidation over a long period of time is thought to strengthen the fear memory (Parsons & Ressler, 2013), whereas reconsolidation blockade immediately after fear memory retrieval can diminish it. In short, decreasing the CR could be achieved via either extinction or reconsolidation blockade, although the latter may be most successful when fear memories are relatively new (Milekic & Alberini, 2002).

To summarize rodent studies of fear conditioning and extinction, the amygdala is a key structure for recognizing salient cues such as potential threats and their predictors, and for mobilizing behavioral fear responses (Davis & Whalen, 2001). Amygdalar output is modulated by the neocortex, especially by the differential projections of the prelimbic and infralimbic regions of the medial prefrontal cortex (Vertes, 2004). In rodents, the prelimbic cortex projects to the lateral amygdala and activity in this region is associated with learning the CS-US predictive pairing during fear conditioning. The infralimbic cortex, in contrast, projects to the inhibitory intercalated neurons and serves to reduce fear after a new CS-noUS association is learned during extinction. When a fear memory is reactivated, it is temporarily labile and vulnerable to manipulation before reconsolidation. The phenomenology and neurocircuitry of fear conditioning and extinction revealed by rodent studies appear to be remarkably well preserved among other species, including humans (Milad, Rauch, Pitman, & Quirk, 2006).

3.1 FEAR CONDITIONING and EXTINCTION IN HEALTHY HUMANS

In human fear conditioning studies, the dependent measure quantifying fear (the UR or CR) is usually a psychophysiological response such as skin conductance response (SCR) or fear-potentiated startle. Functional neuroimaging studies frequently associate SCR with change in brain activity, but less frequently use fear-potentiated startle, as movement disrupts brain imaging. These studies usually use a finger or wrist shock as the US and often compare a conditioned cue (CS+) to an unconditioned or extinguished cue (CS-). Here, we briefly review the methods used in human fear conditioning and extinction studies and describe the basic findings, first in psychophysiological studies and then in functional brain imaging studies.

3.2 PSYCHOPHYSIOLOGICAL METHODS IN HUMANS

In humans, psychophysiological measures of the sympathetic nervous system response to fear or perceived threat most often include SCR, electromyographic (EMG) response, or heart rate response (HRR). SCR is a method of measuring the perspiration-induced electrical conductance or moisture level of the skin, usually on the extremities (i.e. fingers, palms, or feet) and is the most widely used psychophysiological measure in fear conditioning studies.
EMG is a technique that records the electrical activity of skeletal muscles such as the orbicularis oculi (eyelid) or the corrugator supercilli (forehead). Studies of fear-potentiated startle in humans most commonly use EMG to measure eyeblink response or HRR. Heart rates can be measured through various means including pulse oximetry. Electroencephalography (EEG) event-related potential (ERP) studies of conditioning and extinction in healthy humans and PTSD have been reviewed elsewhere (Javanbakht, Liberzon, Amirsadri, Gjini, & Boutros, 2011; Pitman et al., 2012), and will not be discussed here.

3.2.2 PSYCHOPHYSIOLOGICAL FINDINGS IN HUMANS—Healthy humans robustly demonstrate an increase in psychophysiological responses during fear conditioning (e.g., Cohen & Randall, 1984; Hamm & Vaitl, 1996). For example, during fear conditioning acquisition, SCR is higher in response to a CS+ relative to a CS− and is indicative of successful fear learning. Accelerated HRR has also been demonstrated during fear conditioning acquisition, when the aversive stimulus is imminent (Hamm, Greenwald, Bradley, & Lang, 1993). Healthy human studies have also shown CS+ potentiation of the startle reflex, relative to CS-s (Hamm et al., 1993; Lipp, Sheridan, & Siddle, 1994; reviewed in Grillon & Baas, 2003). Furthermore, CS generalization can be demonstrated in healthy humans by utilizing eyeblink as a measure of fear-potentiated startle responses (Lissek et al., 2008).

During extinction, psychophysiological responses to a CS decrease as extinction learning progresses, and during extinction recall, an extinguished CS is associated with significantly lower SCR than an unextinguished CS (Milad, Orr, Pitman, & Rauch, 2005; Milad, Wright, et al., 2007). SCR is thus an inverse measure of extinction recall.

An explicit safety signal is different than extinction because it is a cue that dependably signals that no US will be presented, as opposed to the gradual learning of a CS-noUS association that competes with a CS-US association. However, both involve the inhibition of learned fear. Jovanovic and colleagues have utilized fear-potentiated startle during a conditional discrimination paradigm (AX+/BX−) to assess the role of safety signals learning during fear conditioning (reviewed in Jovanovic, Kazama, Bachevalier, & Davis, 2012). This paradigm was first developed in rats, demonstrating the translational appeal of fear-potentiated startle (Myers & Davis 2004). In the adapted human paradigm, the US is a blast of air to the throat and three different colored shapes (A, B, and X) serve as cues. The paradigm is called AX+/BX− because the US’s pairing with X depends upon the presence of either A or B. A is a threat signal (like a CS) and B becomes a safety signal. In healthy humans, a sudden loud noise during AX potentiates an eyeblink startle response, relative to a loud noise during BX or AB (Jovanovic et al., 2005). During this task, healthy control participants exhibited heightened fear potentiation under the AX condition. During a critical test phase, A is paired with B (without X). Healthy humans show reduced fear-potentiated startle during this AB presentation, indicating successful learning of B as a safety signal and ability to inhibit a learned fear response (Jovanovic et al., 2005).
3.3 BRAIN IMAGING METHODS IN HUMANS

Human fear conditioning and extinction neuroimaging studies have replicated the basic neurocircuitry findings of rodent studies (for review, see Milad & Quirk, 2012; Rauch, Shin, & Phelps, 2006; Sehlmeyer et al., 2009). While human psychophysiological studies measure SCR, HRR, or EMG as the CR, neuroimaging studies measure brain activity. Three-dimensional images of ongoing brain function can be captured via functional magnetic resonance imaging (fMRI), or positron emission tomography (PET). Both methods allow for simultaneous recording of some psychophysiological measures but because fMRI does not require injection of a contrast agent and has the ability to obtain functional and high-resolution structural images in the same session, it is the most commonly used neuroimaging technique. Neither fMRI nor PET can achieve the cellular-level resolution of rodent studies that can utilize surgically implanted electrodes to measure brain activity in a behaving rat. Furthermore, current functional neuroimaging methods are unable to differentiate between excitatory (glutamatergic) and inhibitory (GABAergic) processes (Heeger & Ress, 2002). However, the basic neurocircuitry is assumed to operate similarly within the amygdalae of rodents and humans (see Figure 1 for a schematic of shared amygdala circuitry).

During functional neuroimaging, participants are presented with sounds or pictures as CSs, and finger shocks or other aversive stimuli as USs. Because of the importance of context in fear conditioning and extinction studies, researchers need a way to manipulate this variable (a neuroimaging environment is its own unique context and cannot change during an experimental session). To this end, most groups display images of different settings (e.g., an office room or a library) as a proxy for context, with a discrete CS either superimposed upon (e.g., a red X in the middle of the screen) or embedded within (e.g., an illuminated red lamp in the room) the context.

For example, Milad et al. have developed a two-day conditioning and extinction paradigm in an fMRI scanner to study the brain bases of fear conditioning, extinction, and extinction recall in PTSD (Linnman, Zeffiro, Pitman, & Milad, 2011; Milad et al., 2009; Rougemont-Bücking et al., 2011). The methodology is as follows: Because rodent contextual fear conditioning studies have demonstrated the importance of context pre-exposure (Rudy & O’Reilly, 2001), on day 1, participants are first pre-exposed to images of two different rooms. The rooms are an office and a library, both featuring a lamp. Subsequent conditioning takes place in one of the contexts. During conditioning, two lamp colors (e.g., blue and yellow, the CS+s) both predict finger shock approximately 60% of the time, while another lamp color (e.g., red, the CS−) never predicts shock. After a short break, extinction of one of the CS+s occurs in a different “context” than conditioning. To ensure that participants believe that a US might still be administered, the shock-delivery electrodes remain attached to participants’ fingers for the remaining phases of the experiment. Participants are instructed that they may be shocked, but the US is never again delivered. Importantly, only one of the two CS+s (lamp colors) is presented during extinction learning, allowing subsequent within-subjects comparisons among unextinguished CS, extinguished CS, and the CS−. This comparison occurs the following day, during extinction recall. Such methodology can be used to elucidate brain function during fear conditioning, extinction and extinction recall in healthy humans and in anxiety disorders such as PTSD.
3.3.2 NEUROIMAGING FINDINGS DURING CONDITIONING IN HEALTHY HUMANS—In accordance with the amygdala’s well-established role in fear expression and learning in rodents, many neuroimaging studies of healthy humans have reported amygdala activation during fear conditioning (Alvarez, Biggs, Chen, Pine, & Grillon, 2008; Barrett & Armony, 2009; Buchel, Dolan, Armony, & Friston, 1999; Buchel, Morris, Dolan, & Friston, 1998; Cheng, Knight, Smith, & Helmstetter, 2006; Cheng, Knight, Smith, Stein, & Helmstetter, 2003; Cheng, Richards, & Helmstetter, 2007; Furmark, Fischer, Wik, Larsson, & Fredrikson, 1997; Gottfried & Dolan, 2004; Knight, Nguyen, & Bandettini, 2005; Knight, Smith, Cheng, Stein, & Helmstetter, 2004; LaBar, Gatenby, Gore, LeDoux, & Phelps, 1998; Lang et al., 2009; Linnman, Zeidan, Pitman, & Milad, 2013; Milad, Wright, et al., 2007; Morris & Dolan, 2004; Pine et al., 2001; Tabbert, Stark, Kirsch, & Vaitl, 2006). This amygdala activation co-occurs with increased SCR, which is evidence that amygdala activation is the neural correlate to the CR. The amygdala activates to a CS that has been conditioned to complex USs such as aversive images or video clips (Doronbekov et al., 2005; Klucken et al., 2009) or conditioned to visceral USs such as lower gastrointestinal pain (Kattoor et al., 2013), as well as to a simple unconditioned US such as finger shock (Linnman, Rougemont-Bücking, Beucke, Zeffiro, & Milad, 2011). The amygdala also activates even when the CS is presented outside of conscious awareness (Critchley, Mathias, & Dolan, 2002; Knight, Waters, & Bandettini, 2009), and the resting-state functional connectivity of the human amygdala is altered after fear conditioning (Schultz, Balderston, & Helmstetter, 2012).

For the purposes of comparing human and rodent fear conditioning and extinction studies, the human medial prefrontal cortex (mPFC) can be divided into dorsal and ventral structures. The dorsal mPFC includes the dorsal anterior cingulate cortex (dACC), which is the likely human homologue of the rodent prelimbic cortex (see Figure 1 for a schematic of rat and human mPFC homologues projecting to the same regions of the amygdala, which shares circuitry across species). The ventral medial prefrontal cortex (vmPFC) is the likely homologue of the rodent infralimbic cortex and comprises several structures, such as the medial frontal gyrus (MFG), rostral anterior cingulate cortex (rACC), and subgenual anterior cingulate cortex (sgACC).

Fear conditioning or the expression of conditioned fear in healthy humans has been associated with activity in medial prefrontal brain regions such as dACC and vmPFC (Alvarez et al., 2008; Buchel et al., 1999; Buchel et al., 1998; Klucken et al., 2009; LaBar et al., 1998; Lang et al., 2009; Maier et al., 2012; Marschner, Kalisch, Vervliet, Vansteenwegen, & Buchel, 2008; Milad, Quirk, et al., 2007; Milad, Wright, et al., 2007; Morris & Dolan, 2004; Phelps, Delgado, Nearing, & LeDoux, 2004). SCRs during conditioning are correlated with fMRI response in dACC and dACC thickness (Milad, Quirk, et al., 2007), and with the rate of resting metabolism in dACC (Linnman et al., 2013).

Recent studies have reported hippocampal activation during contextual (Alvarez et al., 2008; Lang et al., 2009; Marschner et al., 2008) and simple cue (Buchel et al., 1999; Knight et al., 2004; Knight et al., 2009) fear conditioning in healthy humans. In a stimulus generalization study, Lissek et al (2013) found that, as stimuli grew increasingly different than a CS+, functional activation in the ventral hippocampus and vmPFC increased. As stimuli grew
increasingly similar to a CS+, functional activation in the insular cortex increased. In addition, several studies have found increased insular cortex activation during fear conditioning in healthy humans (Gottfried & Dolan, 2004; Klucken et al., 2009; Knight et al., 2009; Marschner et al., 2008; Morris & Dolan, 2004; Phelps et al., 2004). The insula also responds to paradigms (e.g., anticipatory anxiety) that elicit more sustained fear responses (Grupe, Oathes, & Nitschke, 2013; Phelps et al. 2001; Somerville, Whalen, & Kelley, 2010). The fear vs. anxiety distinction is discussed more in section 4.1.2.

3.3.3 NEUROIMAGING FINDINGS DURING EXTINCTION IN HEALTHY HUMANS

Several studies of healthy humans have found extinction-associated activation of the amygdala (Gottfried & Dolan, 2004; Milad, Wright, et al., 2007; Phelps et al., 2004) and insula (Gottfried & Dolan, 2004; Phelps et al., 2004). As the CS gradually ceases to predict the US during later stages of extinction learning, activation of the amygdala diminishes. Based upon rodent studies, this likely reflects increased inhibitory influence from the intercalated neurons of the amygdala, which is driven by the vmPFC. Indeed, as the homologue to the rodent infralimbic cortex, the human vmPFC is associated with increased activity during extinction learning (Barrett & Armony, 2009; Gottfried & Dolan, 2004; Kalisch et al., 2006; Linnman et al., 2012; Milad, Wright, et al., 2007) and during extinction recall (Kalisch et al., 2006; Phelps et al., 2004). During extinction recall, vmPFC activation is positively correlated with SCR measures of extinction recall and also with hippocampus activation, indicating a role for these brain regions in the contextual modulation of fear extinction (Milad, Wright, et al., 2007). Recently, Linnman et al. (2012) found that resting-state metabolism in dACC is negatively correlated with extinction recall, and resting-state metabolism in amygdala is positively correlated with functional dACC activation during extinction recall and negatively correlated with functional vmPFC activation during extinction recall.

In summary, functional neuroimaging studies have shown that amygdala, dACC, and insular cortex activity are associated with fear conditioning and the expression of conditioned fear. Furthermore, the hippocampus and vmPFC activate during conditioning in some studies, and may be involved in the processing of contextual information. During extinction learning, the vmPFC is activated. Amygdala activity diminishes as the extinction association is learned. During extinction recall, SCR is negatively correlated with vmPFC and hippocampal activation.

4.1 THE FEAR CONDITIONING MODEL OF PTSD

For decades, psychologists have theorized that pathological anxiety may reflect a failure to extinguish conditioned fear. In the 1970s, a “conditioning model of neurosis” emerged, which held that impaired extinction for certain CRs form the basis of anxieties, phobias, and compulsions (Eysenck, 1979; Pitman & Orr, 1986). In 1980, PTSD became a formal diagnosis in the DSM-III and was considered by many to be uniquely relevant to the conditioning framework because it is the only anxiety disorder to involve an explicit conditioning episode (i.e. the traumatic experience). Of course, the type of US that actually leads to PTSD is much more grievous than the sudden loud noise that Little Albert experienced or the mild finger shock delivered during the fMRI studies just described. And
the conditioned fear response is much more debilitating, sometimes lasting decades after a traumatic experience. However, because the underlying processes appear to be similar, the fear conditioning and extinction framework has proven to be valuable in attempting to understand the mediating neurocircuitry and treatment of PTSD (Holmes & Singewald, 2013).

PTSD is a severe anxiety disorder that affects approximately 10-30% of individuals who have experienced or witnessed a psychologically traumatic event (de Vries & Olff, 2009; Dohrenwend et al., 2006; Weiss et al., 1992) such as sexual assault, combat, natural disaster, motor vehicle accident, or witnessing the death or serious injury of another individual. Individuals with PTSD experience three primary symptoms: 1) unwanted re-experiencing of the event, often in the form of intrusive memories and nightmares; 2) avoidance of reminders of the traumatic event, even in the absence of real danger; and 3) ongoing hypervigilance, including trouble with sleep and concentration. Reminders of the traumatic experience are particularly distracting and distressing to individuals with PTSD (Hayes, VanElzakker, & Shin, 2012). While avoidance and hypervigilance can occur in the absence of trauma-related cues, re-experiencing events may be triggered by a previously neutral cue that was conditioned to the traumatic experience.

PTSD is frequently conceptualized as a memory disorder, within a Pavlovian fear conditioning and extinction framework (Elzinga & Bremner, 2002; Jovanovic & Ressler, 2010; Rubin, Berntsen, & Bohni, 2008; Shin & Handwerger, 2009). According to this framework, multiple cues become associated with a traumatic experience, forming strong sensory memories. This can occur in a manner that Pavlov demonstrated by simultaneously conditioning to a buzzer, a metronome, and tactile stimulation. Or it can happen by stimulus generalization, like when Little Albert became wary of objects that shared some trait with a white rat (the CS) whose presence predicted a loud, startling noise (Harris, 2011; Watson & Rayner, 2000). Multiple cues can be conditioned to the same US and as we will review below (section 5.1), PTSD is associated with increased acquisition of fear conditioning. Furthermore, anxiety disorders are associated with increased propensity for CS generalization (Lissek et al. 2005). Therefore, it is likely that multiple cues would be capable of triggering a conditioned fear response in an individual with PTSD. Upon encountering conditioned reminders, individuals with PTSD re-live their traumatic experience through sensory memories. These memories are often disjointed and incomplete, with a disorganized and sometimes out-of-sequence narrative (Brewin, 2011). This may reflect the diminished sensory representations of the subcortical pathway discussed earlier (section 2.1; see also Pitman et al., 2012).

If PTSD symptoms emerge from a traumatically stressful conditioning episode, the failure to extinguish or maintain extinction of that conditioned fear allows symptoms to persist. For example, after the invasion of Iraq, US troops frequently faced improvised explosive devices (IEDs) hidden in trash piles along the side of the road. If a soldier survives an IED attack, this event forms a powerful memory and powerful associations. Ideally, such a veteran who has returned to the United States should learn, upon several exposures to trash piles along his street, that they no longer predict violence. If this extinction fails to happen, then even years after the danger has passed, a previously neutral cue (trash, the CS) that was associated
with a grave danger (IED explosions, the US) may trigger a fear response (CR) and painful memories. Indeed, even internal cues such as autonomic arousal or thirst can become CSs. For a veteran of war in Iraq, the CSs that cause a fear response could be the sight of a trash pile, the presence of a crowd of people, the feel of hot temperatures under bright sunlight, the sound of Arabic being spoken, the smell of diesel fuel from military vehicles, internal states such as feelings of frustration, or any other cue that was present during the explosion. In his studies of multiple CSs, Pavlov reported that extinguishing one CS also extinguished the others, although subsequent research has found that this is not the case and that extinction is cue-specific (Myers & Davis, 2007).

Although humans are only able to attend to approximately four discreet cues at a time (e.g., Bettencourt & Somers, 2009), our memory of experiences normally feels rich and complete. For example, if you were asked where you were during breakfast this morning, you could easily think of more than four details: the tastes and smells of the food, the table and chair where you sat, what you were wearing, what you ate, what you read or talked about while eating, and even how you felt are all details that you did not plan on remembering, but were encoded into one episodic memory (Rudy, 2009; Rudy & Tyler, 2010). Pavlov referred to this as “dynamic stereotype”—a neurological mapping of the environment. This is what we mean when we refer to the “context” of a memory, including fear memories. The forming of such cognitive representations of context is rapid and automatic. Context is a crucial variable that serves as something more than just another CS during fear conditioning—it also acts as a safety signal after extinction learning (reviewed in Bouton, Westbrook, et al., 2006; Maren et al., 2013; Rudy, 2009). This means that extinction is particularly influenced by the context in which it occurs (Bouton, García-Gutiérrez, Zilski, & Moody, 2006).

For example, after a soldier comes back to the United States and repeatedly sees innocuous trash piles, the association between trash and IEDs would ideally become extinguished, at least in the context of the United States. However, if that soldier were re-deployed to Iraq, the association between roadside trash and danger might renew, because that association was appropriate in the context of occupied Iraq, and extinction learning did not occur in that context. Thus, the process of fear extinction has been described as a context-modulated form of inhibitory learning (Bouton, 2004). In other words, because extinction learning forms a new association (CS-noUS) as opposed to “breaking” the conditioned CS-US association, extinction recall involves accessing competing CS associations (CS-US vs. CS-noUS). Therefore, if extinction learning takes place in a different context than conditioning, the CR becomes context-dependent. Just as an ambiguous word (e.g. “lead”) is given meaning by context (i.e. “Metal weights are made of lead” vs. “The winner of the race was in the lead from the start”), the ambiguous CS (a pile of trash) is given meaning by the context (suburban United States vs. occupied Iraq).

However, context often doesn’t serve as an effective safety signal for individuals with PTSD. Even while in the United States, and even after seeing innocuous trash piles many times, a veteran with PTSD may fail to extinguish the association between trash piles and IED explosions (CS-US). The CS continues to elicit a CR, namely distressing sensory memories of an explosion. Empirical fear conditioning and extinction studies of this
phenomenon will be discussed in the next section (5.1). Here, we relate renewal, spontaneous recovery, and reinstatement to the clinical presentation of PTSD.

Exposure-based therapy is used in the treatment of PTSD and is conceptually based upon fear extinction. The phenomenon of renewal (return of a CR in the conditioning context when extinction took place in a different context) is highly relevant to this type of therapy. Measured exposure to progressively stronger trauma-related stimuli and memories (CSs) serves to create a new association between those cues and safety (CS-noUS; Cukor, Spitalnick, Difede, Rizzo, & Rothbaum, 2009; Rothbaum & Davis, 2003). An individual may spend hours in a therapist's office, talking about his traumatic experience in great detail, effectively extinguishing the CS-US association for any cues that remind him of the trauma. However, because extinction is context-specific, the CS-noUS association arising from exposure therapy only exists in his therapist's office and not in other contexts. In other words, if fear conditioning took place in Iraq and fear extinction takes place in a therapist's office, when an individual with PTSD is confronted with a now-ambiguous CS (e.g., trash) in a third context (e.g., a suburban street), the CS-US (trauma) association will beat out the CS-noUS (therapy) association. The CS-noUS fails to oppose the stronger, more salient CS-US association. This is the reason that therapists conduct sessions in multiple settings and use virtual reality to mimic multiple contexts (e.g. Difede et al., 2007; Gerardi, Rothbaum, Ressler, Heekin, & Rizzo, 2008; Rothbaum et al., 1999). Extinction therapy and pharmacological enhancers of extinction will be discussed in more detail below (section 6.2).

Delayed onset PTSD, in which the emergence of symptoms does not occur until six months or more post-trauma (APA, 2000), occurs in 8-25% of individuals with PTSD (Bryant, O'Donnell, Creamer, McFarlane, & Silove, 2013; Frueh, Grubaugh, Yeager, & Magruder, 2009), and may be the clinical manifestation of the phenomenon of spontaneous recovery (the return of a CR after a long delay since its extinction). In contrast, reinstatement of the CR occurs when a US is re-experienced, even when it is not paired with the CS. This can manifest in PTSD patients who have a sort of “relapse” of symptoms after successful therapy upon experiencing another stressful event. For example, Pitman (2011) described a World War II veteran who had experienced nightmares and flashbacks for one year after the war's end. He was then symptom-free for 30 years, until experiencing another highly stressful event (i.e., another US). The night after receiving a terminal cancer diagnosis, he experienced nightmares again—but about combat, not cancer.

Fear conditioning paradigms provide a useful framework for studying the canonical PTSD symptom of re-experiencing the trauma through powerful but disjointed memories upon exposure to any reminder. Although extinction failure can explain the persistence of trauma-related memories and distress in PTSD, reconsolidation also likely contributes. Repeated reactivation of traumatic memories triggered by CSs can reconsolidate and strengthen them (Inda, Muravieva & Alberini, 2011; Parsons & Ressler, 2013). However, the fear conditioning model may not explain other PTSD symptoms such as anxiety in the absence of trauma reminders (unless stimulus generalization has occurred), or more complex cognitive and biological symptoms such as dissociation or endocrine abnormalities (reviewed in Zoladz & Diamond, 2013).
4.1.2 FEAR VS. ANXIETY

Pavlovian fear conditioning paradigms that use a cue as a CS effectively provoke the fear response, which is a phasic (discrete) response to a specific stimulus or event (e.g., a tone that predicts a shock). However, this paradigm might not appropriately model the ongoing anxiety in PTSD that leads to excessive hypervigilance and avoidance of trauma reminders. Such anxiety has been conceptualized as “sustained fear” and can instead be investigated using paradigms that put individuals into a general state of uncertainty, such as contextual (as opposed to cued) fear conditioning or paradigms that emphasize unpredictability (Davis, Walker, Miles, & Grillon, 2010). As detailed in section 2.1, during contextual fear conditioning one learns to associate an aversive event (US) with a particular context rather than a distinct cue. The use of cue conditioning vs. contextual conditioning paradigms has provided insight into the neural basis of the overlap and distinction between fear (discussed in more detail in earlier sections) and anxiety. Both involve the amygdala, which responds to stimuli in the environment that predict biologically relevant events (Davis & Whalen 2001). Autonomic and behavioral responses to both types of paradigms are generated via projections from the amygdala to hypothalamic and brainstem targets (Davis et al., 2010). Research suggests that contextual fear learning is dependent on both the amygdala and hippocampus (Kim & Fanselow, 1992; Phillips & LeDoux, 1992). The sustained fear responses elicited by these paradigms rely on the bed nucleus of the stria terminalis (BNST; reviewed in Davis et al. 2010). The insula is also heavily interconnected with both the BNST and amygdala, and is involved in producing sustained fear responses such as anticipatory anxiety (Grupe et al., 2013; Phelps et al., 2001; Somerville et al., 2010).

5.1 CONDITIONING AND EXTINCTION IN PTSD

5.1.2 PSYCHOPHYSIOLOGICAL STUDIES

As detailed in section 3.2.2, psychophysiological measures are a valuable tool for assessing fear conditioning and extinction in healthy humans. Researchers have also used these tools to explore fear conditioning and extinction abnormalities in PTSD (reviewed in Lissek et al., 2005; Pitman et al, 2012). Based upon that literature, Guthrie and Bryant (2006) suggested that alterations in psychophysiological responses that characterize PTSD during fear conditioning and extinction include 1) greater psychophysiological response to USs; 2) greater psychophysiological responses during conditioning; and 3) higher continued psychophysiological responses to CSs during extinction learning. We will focus on the latter two points because the first is not as directly relevant to fear conditioning and extinction.

During the acquisition conditioned fear, participants with PTSD exhibit larger differential (CS+ vs. CS−) SCR, EMG, and HRR responses relative to individuals without PTSD (Blechert, Michael, Vriends, Margraf, & Wilhelm, 2007; Glover et al., 2011; Lissek et al., 2005; Orr et al., 2000; Peri, Ben-Shakhar, Orr, & Shalev, 2000), although this increased response during acquisition has not been replicated in all studies (Glover et al., 2011; Milad et al., 2008; Milad et al., 2009).

Studies using SCR, EMG, and HRR also show heightened CRs in individuals with PTSD during extinction learning (Blechert et al., 2007; Lissek et al., 2005; Orr et al., 2000; Peri et
al., 2000) and next-day extinction recall (Milad et al., 2008; Milad et al., 2009), consistent with the idea that people suffering from PTSD have difficulty learning and remembering that stimuli that used to predict threat no longer do. Research using monozygotic twins discordant for combat exposure suggests that this extinction deficit may be an acquired characteristic of PTSD (Milad et al., 2008); however, reduced extinction learning before trauma exposure has also been linked to PTSD symptom severity after trauma exposure, which suggests that impaired extinction may be a vulnerability factor for PTSD (Guthrie & Bryant, 2006).

Fear-potentiated startle paradigms have also proven to be a useful tool for exploring PTSD-related abnormalities in fear learning and memory. For example, individuals with PTSD show elevated fear-potentiated startle responses in “threat” contexts (Grillon, Morgan, Davis, & Southwick, 1998; Morgan, Grillon, Southwick, Davis, & Charney, 1995). Fear-potentiated startle can also be used to quantify fear during fear conditioning and extinction paradigms. Such studies demonstrate that while individuals with PTSD are being conditioned, they exhibit enhanced fear-potentiated startle during the presentation of both CS+ and CS− during conditioning (Grillon & Morgan, 1999; Glover et al., 2011; Norrholm et al. 2011) and sustained fear responses to the CS+ during extinction (Norrholm et al. 2011). Interestingly, heightened fear-potentiated startle to the CS+ predicted re-experiencing symptoms (Glover et al., 2011; Norrholm et al. 2011) while heightened fear-potentiated startle to the CS− predicted hyperarousal symptoms (Glover et al., 2011). One study also found that an attentional bias towards threat sometimes reported in PTSD (reviewed in Hayes et al., 2012) was associated with increased fear-potentiated startle response to both CS+ and CS− during fear conditioning and extinction (Fani et al., 2012).

One explanation for the fact that individuals with PTSD show elevated fear responses to both CS+ and CS− is that they have an impaired ability to translate the presence of a safety signal into inhibited fear responses (Jovanovic at al., 2012). Using the AX+/BX− conditional discrimination paradigm described in section 3.2.2 (Jovanovic et al., 2005), researchers learned that despite self-reports indicating that individuals with PTSD appropriately learn that a particular cue (B) signals safety, they are still unable to inhibit their fear-potentiated startle response in its presence (Jovanovic et al., 2009). This manifested in an association between heightened fear-potentiated startle during the BX and AB trials and PTSD symptom severity. This failure of individuals with PTSD to learn the “B” safety signal has been replicated, and in a subsequent study, heightened fear-potentiated startle reactivity to AB trials was associated with hyperarousal symptom severity (Jovanovic et al., 2010). These results, along with psychophysiological studies demonstrating impaired fear extinction, are consistent with a general failure to inhibit learned fear, and may be an intermediate phenotype between the clinical symptoms of PTSD and the dysfunctional neurocircuitry found in this disorder (Jovanovic & Norrholm, 2011).

5.1.3 NEUROIMAGING STUDIES OF PTSD

Psychophysiological abnormalities in conditioning and extinction in PTSD are also reflected in neuroimaging studies. Compared to healthy controls, individuals with PTSD show hyper-responsivity of brain regions associated with fear expression and conditioning, but hypo-
responsivity of brain regions associated with fear extinction and extinction recall. Imaging studies have revealed abnormalities in PTSD at each stage of the fear conditioning and extinction paradigm. Essentially, in PTSD the fear “accelerators” (amygdala and dACC) fail to turn off while the “brakes” (vmPFC and hippocampus) fail to turn on. These are the brain correlates to the increased psychophysiological responses during conditioning and extinction seen in PTSD. In the first neuroimaging study of fear conditioning and extinction in PTSD, Bremner and colleagues (2005) found that relative to healthy, non-abused female controls, women with childhood sexual-abuse-related PTSD showed greater left amygdala and dACC activation in the contrast between a fear acquisition condition (in which a shock to the forearm was predicted by a picture of a blue square) and a control condition (in which participants received random shocks with no paired CS). During extinction, the PTSD group had relatively less activation in rACC and subgenual cortex. This study did not investigate extinction recall or manipulate context.

As described earlier, Milad and colleagues have used a two-day conditioning and extinction paradigm that manipulates context. Using this paradigm in an fMRI environment, this group has found evidence that, in individuals with PTSD, structures that are involved in fear expression over-activate while structures involved in fear inhibition under-activate or deactivate (Linnman, Zeffiro, et al., 2011; Milad et al., 2009; Rougemont-Bücking et al., 2011). Relative to the non-PTSD control group, the PTSD group had greater amygdala activation to the US (a finger shock; Linnman, Zeffiro, et al., 2011). During late conditioning on day 1, they found increased dACC activation in PTSD, relative to the trauma-exposed non-PTSD control group (Rougemont-Bücking et al., 2011). During late extinction on day 1, in a contrast of the extinguished CS+ versus the never-conditioned CS−, the PTSD group had greater amygdala activation, suggesting that individuals with PTSD fail to learn fear extinction to the extent of trauma-exposed controls (Milad et al., 2009). On the following day, fear extinction recall was tested. The PTSD group exhibited relatively greater dACC activation and relatively less activation in the hippocampus and vmPFC during extinction recall. This implies that the regions responsible for identifying the ambiguous CS as safe (hippocampus and vmPFC) failed to engage and the region responsible for identifying learned fear associations (dACC) remained engaged. Thus, in PTSD, the hippocampus may fail to differentiate safe from unsafe contexts after extinction learning while the vmPFC fails to retain the extinction of learned fear responses in the non-traumatic context.

5.2 NEUROCIRCUITRY MODEL OF PTSD

Interestingly, the finding of hyper-responsivity of the amygdala and dACC and hypo-responsivity of vmPFC structures generalizes beyond studies of fear conditioning and extinction (VanElzakker, Staples, & Shin, in press). For example, the amygdala is hyperresponsive in PTSD relative to control groups during trauma-related (Shin et al., 2004), trauma-unrelated emotional (Brohawn, Offringa, Pfaff, Hughes, & Shin, 2010) and emotionally neutral (Bryant et al., 2005) tasks. Likewise, the dACC is hyperresponsive relative to controls during trauma-related (Fonzo et al., 2013), trauma-unrelated emotional (Hayes, LaBar, Petty, McCarthy, & Morey, 2009) and emotionally neutral (Shin et al., 2007, 2011) tasks.
Similarly, functional abnormalities in other fear conditioning and extinction-related brain structures point to a general neurocircuitry dysfunction in PTSD. Several studies using other tasks have reported significant negative correlations between vmPFC activation and PTSD symptom severity (reviewed in Hughes & Shin, 2011). This is evidence that reduced activation of vmPFC in PTSD is not specific to fear extinction learning and recall, but rather appears to be a more general functional property of that brain region. In PTSD, the vmPFC also shows diminished activation or even deactivation relative to controls in response to reminders of trauma (e.g., Lanius et al., 2001; Shin et al., 1999), in tasks using emotional but trauma-unrelated stimuli such as facial expressions (e.g., Offringa et al., 2013; Williams et al., 2006), and even during emotionally neutral tasks (e.g., Moores et al., 2008).

Functional imaging studies of the hippocampus are mixed, with some studies showing increased and some showing decreased activation (Hughes & Shin, 2011). Type of task, type of stimuli (e.g., trauma-related or -unrelated), and type of analysis may all be important factors in these inconsistencies. However, individuals with PTSD consistently have smaller hippocampi than control groups (Woon, Sood, & Hedges, 2010). This may be related to reduced processing power for appropriate differentiation between trauma context and safe contexts. Furthermore, given its role in contextual, episodic, emotional, spatial, and declarative memory processes (LaBar & Cabeza, 2006; Rudy, 2009; Rudy et al., 2004; Squire & Zola-Morgan, 1991), suboptimal hippocampus function may be related to the deficits in these processes seen in PTSD (Astur et al., 2006; Bremner et al., 1999; Bremner et al., 2003; Brohawn et al., 2010; Hayes et al., 2011; Shin & Liberzon, 2010).

Because PTSD is defined by a conditioning event, one may assume that these abnormalities emerge during the trauma or as PTSD emerges. However, another possibility is that this neurocircuitry represents a familial vulnerability factor that interacts with a traumatic experience to manifest PTSD symptoms. For example, a study of monozygotic twins discordant for combat trauma suggested that a small hippocampus is a familial vulnerability factor for PTSD (Gilbertson et al., 2002). Another recent twin study using PET found that veterans with PTSD and their identical co-twins had higher resting-state regional cerebral metabolic rates for glucose (rCMRglu) in the dACC relative to the veterans without PTSD and their co-twins (Shin et al., 2009). Furthermore, rCMRglu in the combat-unexposed twins was positively correlated with PTSD symptom severity in their combat-exposed identical twins. This is evidence that resting-state dACC hyperactivity is a familial risk factor for PTSD that may predate trauma exposure or PTSD. Twin and genetic studies have provided additional evidence that genes may explain some of the relationship between PTSD and abnormalities in fear conditioning and extinction.

5.3 PTSD-RELATED GENES AND FEAR EXTINCTION

A brain-based familial vulnerability to PTSD is likely to be at least partially explained by genotype. Indeed, a study comparing SCRs of healthy monozygotic to healthy dizygotic twins demonstrated a moderate genetic contribution to fear conditioning and fear extinction (Hettema, Annas, Neale, Kendler, & Fredrikson, 2003). Several PTSD-related genes have also been linked to abnormal fear conditioning and extinction in healthy humans and rodents, providing indirect evidence that the abnormalities in PTSD revealed by the fear
conditioning model may represent a vulnerability endophenotype (Amstadter, Nugent, & Koenen, 2009; Lonsdorf & Kalisch, 2011). PTSD-associated genes that are also implicated in facilitated fear conditioning or attenuated extinction include the serotonin transporter gene, dopamine-related genes, and neuroplasticity-related genes (other genes are certainly involved; the genetics of PTSD are reviewed in Cornelis, Nugent, Amstadter, & Koenen, 2010; Digangi, Guffanti, McLaughlin, & Koenen, 2013; Pitman et al., 2012; Skelton, Ressler, Norrholm, Jovanovic, & Bradley-Davino, 2011 and the genetics of fear conditioning and extinction are reviewed in Johnson et al., 2012; Lonsdorf & Kalisch, 2011).

Genetic studies have found associations between PTSD and a short (s) allele polymorphism that leads to low expression of the serotonin transporter gene (5-HTTLPR, also known as 5-HTT, SERT, or SLC6A4). A study of healthy individuals by Lonsdorf et al. (2009) used the fear-potentiated startle paradigm and found that 5-HTTLPR s-s and s-l allele carriers had larger potentiated conditioned startle responses, relative to l-l allele carriers. Similarly, Garpenstrand, Annas, Ekblom, Oreland, and Fredrikson (2001) used SCR and also found enhanced conditioning in individuals with at least one 5-HTTLPR s-allele. Another study investigated the social learning of fear conditioning (Cryan et al., 2009). They found enhanced SCR reactions to a CS in s-allele carriers who had observed another person undergo fear conditioning, even though those participants never experienced a CS-US association themselves (Cryan et al., 2009). Previous studies have found associations between 5-HTTLPR and TPH2 (tryptophan hydroxylase-2, an enzyme that influences presynaptic serotonin synthesis) “risk” genotype with PTSD and major depression (Bryant et al., 2010; Goenjian et al., 2012). Hermann et al. (2012) used fear conditioning and extinction within an fMRI environment to study the influence of 5-HTTLPR and TPH2 genotype in healthy individuals. They found that 5-HTTLPR risk genotype predicted increased activation in right insula to the US and to the CS during late conditioning. 5-HTTLPR risk genotype interacted with childhood adversity to predict increased amygdala activation during extinction learning. TPH2 risk genotype interacted with childhood adversity to predict greater amygdala activation during conditioning and greater vmPFC activation during extinction. Finally, a combination of 5-HTTLPR and TPH2 risk genotypes predicted increased dACC activation during extinction learning. Another recent study of healthy individuals examined the effect of 5-HTTLPR and neuropeptide S receptor (NPSR1) genotype on contextual fear conditioning and extinction in a virtual reality paradigm (Glotzbach-Schoon et al., 2013). They found that individuals with risk alleles in both genes exhibited higher fear-potentiated startle responses in the conditioned context relative to the “safe” context. NPSR1 genotype was also associated with higher self-reported anxiety levels in the conditioned context, relative to the safe context. They did not find genotype differences during extinction. These studies show a pattern of facilitated fear conditioning that carries over into extinction in 5-HTTLPR risk allele carriers (reviewed in Lonsdorf & Kalisch, 2011).

Several studies have found abnormal fear conditioning and extinction associated with dopamine-related genes. For example, Raczkka et al. (2011) found that polymorphism in the dopamine active transporter 1 gene DAT1 predicted faster extinction learning in healthy
humans. In a study of 60 healthy individuals, Huertas et al. (2010) found that a polymorphism in the DRD2 dopamine receptor D2 gene predicted significantly attenuated SCR during early conditioning. The gene responsible for the enzyme catechol-O-methyltransferase (COMT) is related to synaptic availability of catecholamines such as dopamine and norepinephrine, especially in the prefrontal cortex. COMT genotype predicted deficits in extinction learning in healthy humans (Lonsdorf et al., 2009). A recent study showed that a COMT risk allele predicted higher fear-potentiated startle regardless of PTSD diagnosis (Norrholm, et al. 2013). That same study found that in individuals with PTSD, the risk allele genotype predicted impaired fear inhibition to a CS— and was associated with DNA methylation status on four loci, two of which were associated with the impaired fear inhibition.

Studies examining the relationship between the neuroplasticity-related brain-derived neurotropic factor (BDNF) gene and fear conditioning and extinction in healthy humans have been somewhat mixed. Most BDNF studies have focused on one single-nucleotide polymorphism (SNP): the Val66Met SNP, which substitutes a valine (Val) for methionine (Met) at codon 66. Hajcak et al. (2009) examined this SNP and found a significant genotype × stimulus interaction in healthy individuals undergoing a fear-potentiated startle conditioning paradigm. Homozygous carriers of the Val allele had a significantly higher fear-potentiated startle response to a CS than Met-carriers. Similarly, Lonsdorf et al. (2010) found that among healthy individuals undergoing fear conditioning and extinction, Val-carriers showed increased conditioned fear-potentiated startle responses during late conditioning and early extinction. In contrast to those two studies, Soliman et al. (2010) found that the Met allele predicted impaired extinction learning in mice and in healthy humans. Furthermore, in their study the Met allele was associated with decreased vmPFC activation and increased amygdala activation during extinction. Another study reported inconsistent data in that homozygous Val-carriers showed heightened SCR to both conditioned and unconditioned cues during reconditioning of a previously extinguished association (Raczka et al., 2011). Given the role of BDNF in hippocampal-dependent processes (Tyler, Alonso, Bramham, & Pozzo-Miller, 2002) as well as fear and anxiety (Frielingdorf et al., 2010), future research may emphasize contextual fear conditioning to better elucidate the role of the Val66Met SNP of BDNF. Other neuroplasticity-related genes may also be prime targets for PTSD-related fear conditioning and extinction research. For example, in mice, the statmin gene, an inhibitor of microtubule formation, and the Grp gene, which encodes gastrin-releasing peptide, are highly expressed in the lateral nucleus of the amygdala and appear to be necessary for fear conditioning (Shumyatsky et al., 2005). Their human homologues may be attractive targets for pharmacological targeting in PTSD.

In summary, serotonin, dopamine, and neuroplasticity-related genes are among those that explain variance in fear conditioning and extinction, and are also related to PTSD. Research in animal models using genome-wide approaches may provide future gene targets in PTSD and other anxiety disorders (Sokolowska & Hovatta, 2013). Interestingly, one-third of the variance in the vulnerability to anxiety disorders and one-third of the variance in human fear conditioning can be attributed to genetics (Lonsdorf & Kalisch, 2011). This raises the

Neurobiol Learn Mem. Author manuscript; available in PMC 2015 September 01.
possibility that abnormalities in fear conditioning and extinction may be a general vulnerability factor across multiple anxiety disorders.

6.1 EXTINCTION IN OTHER ANXIETY DISORDERS

The National Institute of Mental Health (NIMH) recently issued a strategic plan that calls for research using dimensions of observable behavior and neurobiological measures to characterize mental disorders (Research Domain Criteria [RDoC]; Insel et al., 2010; Morris & Cuthbert, 2012). This strategic plan emphasizes examining a single domain (such as fear conditioning/extinction) at multiple levels (from genes to brain function to observable behavior) that cuts across traditional disorder categories. Consistent with this strategic plan, researchers have begun to examine fear conditioning and extinction measures across diagnostic categories (e.g., Holt et al., 2009; Lissek et al., 2005). Here, we focus on studies of anxiety disorders other than PTSD. In general, these studies have included psychophysiological (e.g., SCR, EMG) and self-report measures and have generally shown resistance to extinction in anxiety disorders. For example, in the first study of conditioning in psychopathology, Pitman & Orr (1986) examined SCRs during differential conditioning and extinction in a mixed anxiety disorder cohort, consisting mostly of individuals with generalized anxiety, panic disorder, and social phobia. During the extinction phase, anxiety disorder patients (for whom an angry facial expression had been the CS+) exhibited larger SCRs to the CS+ versus CS−, whereas the control group did not show this difference. This suggests that conditioned fear responses are resistant to extinction in anxiety disorders, although this study alone could not shed light on potential differences between the anxiety disorders.

Focusing on social phobia, Hermann, Ziegler, Birbaumer, and Flor (2002) studied self-report and psychophysiological correlates of conditioning and extinction in men with generalized social phobia and healthy men. Unlike more typical Pavlovian conditioning studies, the CSs in this study were neutral facial expressions and the US was an aversive odor. The social phobia group did not differ from the healthy comparison group on measures of differential (CS+ vs. CS−) conditioning, although they did show a US expectancy bias (i.e., they expected both the CS+ and CS− to be followed by the US). In addition, although the control group showed a decline in differential skin conductance responses during extinction, the social phobia group maintained their differential SCR throughout extinction. In other words, the generalized social phobia group was more resistant to extinction in terms of SCR. Whether groups with more circumscribed social phobia (e.g., performance-only subtype) would show similar abnormalities is not clear. An fMRI study of social phobia revealed no significant between-group brain activation effects during extinction (Schneider et al., 1999); however, SCR was not measured in that study.

Resistance to extinction has also been reported in panic disorder. In a more typical paradigm involving neutral photographs as the CSs and shock as the US, Michael, Blechert, Vriends, Margraf, and Wilhelm (2007) studied differential fear conditioning in patients with panic disorder and healthy control participants. The two groups did not differ on conditioned responses during acquisition. However, during extinction, the panic disorder group exhibited larger SCRs to the CS+ and evaluated the CS+ more negatively than the control group.
contrast, using a simple conditioning procedure, Del-Ben and colleagues (2001) failed to find extinction abnormalities in individuals with panic disorder.

The results of one study suggest that resistance to extinction can also occur in children with anxiety disorders. Liberman and colleagues (2006) included such children (with social phobia, generalized anxiety disorder, separation anxiety, or specific phobia) and non-anxious children in a conditioning paradigm involving picture CSs and a loud tone US (Liberman, Lipp, Spence, & March, 2006). In contrast to the non-anxious children, the anxious children rated the CS+ as more fear provoking than the CS− after extinction. In addition, startle blink magnitude during the extinction phase was larger for the CS+ versus CS− in anxious children, but not in non-anxious children.

Most previous studies of extinction in anxiety disorders have focused on extinction learning rather than on the recall of extinction memories. In a study of the latter type, Milad et al., (2013) used a validated Pavlovian fear conditioning and extinction paradigm (described in section 3.3) to examine individuals with obsessive-compulsive disorder (OCD) and healthy controls. Functional MRI and SCR data were gathered during conditioning and extinction (on day 1) and during extinction recall (on day 2). Although the OCD and healthy control group exhibited a similar degree of conditioning and extinction learning, as determined by SCR measures, the OCD group had lower extinction recall index scores than the control group. In other words, the OCD group had more difficulty recalling extinction training compared to the healthy control group. In addition, during extinction recall, activation in the vmPFC, posterior cingulate cortex, cerebellum, and putamen was lower in the OCD vs. control group. Furthermore, vmPFC activation during extinction training was also lower in OCD. Contrary to prediction, within the OCD group, OCD symptom severity scores were positively correlated with vmPFC, cerebellum, and posterior cingulate activation during extinction recall. OCD symptom severity scores were negatively correlated with activations in the dACC and insular cortex in this same condition. Although OCD appears to share some extinction abnormalities with PTSD (e.g., impairment in extinction recall and diminished activation in vmPFC during extinction recall), OCD is also marked by other abnormalities that were not observed in PTSD (e.g., less activation in the posterior cingulate cortex during extinction recall). The unexpected relationship between OCD symptom severity and activation during extinction recall may indicate the presence of other factors that contribute to extinction abnormalities in OCD.

The findings above are generally consistent with the idea that conditioned fear is resistant to extinction (during extinction learning and when tested after a delay) in anxiety disorders. Future studies of extinction should directly compare different anxiety disorder groups within the same study. Minimal extinction differences between specific anxiety disorder groups could potentially suggest that they lie on the same underlying continuum and may have a common pathophysiology. Thus far, few studies have examined the neural correlates of extinction impairment across anxiety disorders. Nevertheless, the psychophysiological evidence of impaired extinction in anxiety disorders might suggest similar functional abnormalities in the amygdala, dACC, vmPFC, and hippocampus among these disorders. Indeed, neuroimaging studies involving other paradigms have revealed functional abnormalities in these very structures across anxiety disorders, although the direction of the
vmPFC abnormality appears to differ between PTSD and other anxiety disorders (e.g., Etkin & Wager, 2007) and the pathophysiology of OCD appears to involve thalamo-corticostriatal circuits as well (Shin & Liberzon, 2010). Additional neuroimaging studies of extinction across anxiety disorders are needed to confirm speculations of shared pathophysiology.

6.2 DRUG AND NEUROMODULATION TREATMENTS FOR PTSD AND OTHER ANXIETY DISORDERS

As we described earlier, extinction involves repeatedly presenting the CS+ (e.g., tone) without the US (e.g., shock) such that the individual learns that the CS+ no longer predicts the US, and the CR (e.g., freezing in rodents or SCR in humans) diminishes. Repeated exposure to trauma-related cues lies at the heart of exposure therapy and related cognitive-behavioral techniques, which are the most well accepted treatments for PTSD and other anxiety disorders (e.g., Hofmann, Sawyer, Korte, & Smits, 2009; Hofmann & Smits, 2008).

Studies involving the pharmacologic manipulation of extinction in rodents and in healthy humans have helped identify new treatment options for individuals with anxiety disorders. D-cycloserine (DCS) is one of the most fruitful examples of this (Davis, 2011; Hofmann, Wu, & Boettcher, 2013). DCS is an antibiotic that also acts as a partial agonist at the glycine site of the NMDA glutamate receptor and has been shown to facilitate extinction learning in rodents (Davis, Ressler, Rothbaum, & Richardson, 2006; Norberg, Krystal, & Tolin, 2008). In clinical studies, DCS is typically administered prior to exposure sessions and has been found to facilitate symptomatic improvement in phobia (e.g., Ressler et al., 2004), OCD (e.g., Chasson et al., 2010; Wilhelm et al., 2008), social anxiety disorder (Hofmann, Smits, et al., 2013), and panic disorder (Otto et al., 2010), although it may be less effective in PTSD (de Kleine, Hendriks, Kusters, Broekman, & van Minnen, 2012; Litz et al., 2012).

DCS appears to be most effective in treatment sessions during which there is within-session decline in fear (Hofmann, Wu, et al., 2013; Smits et al., 2013), and perhaps smaller such declines could explain the relatively lackluster findings in PTSD thus far. Although recent studies have suggested that DCS may expedite improvement, rather than improve overall outcome (Chasson et al., 2010; Hofmann, Smits, et al., 2013; Norberg et al., 2008), DCS could still be clinically useful by providing faster symptom relief, decreasing the risk of early termination of treatment, and decreasing the cost of care.

Other agents have been shown to improve extinction and may thus prove helpful in the treatment of anxiety disorders. Some are already widely utilized in the treatment of anxiety disorders (e.g., serotonin reuptake inhibitors; SRIs) and others are not (e.g., glucocorticoids, cannabinoids). With regard to SRIs, when administered before extinction training in rats, venlafaxine facilitated between-session extinction (Yang et al., 2012). In addition, chronic administration (21 days) of venlafaxine prevented reinstatement of conditioned fear (Yang et al., 2012). When administered after extinction in rats, fluoxetine prevented stress-induced reinstatement of conditioned fear (Deschaux et al., 2013). When combined with extinction training in mice, fluoxetine facilitated extinction (Karpova et al., 2011; but see also Burghardt, Sigurdsson, Gorman, McEwen, & LeDoux, 2013, which suggests that chronic treatment with citalopram before extinction training impairs extinction in rats). Recently, researchers have investigated whether pre-treatment with a SRI before conditioning/
extinction training can facilitate extinction in healthy humans. In a randomized, placebo-controlled trial, Bui et al. (2013) pre-treated 18 participants with escitalopram (10 mg/day) and 20 participants with placebo for 14 days. At the end of this treatment period, participants underwent Pavlovian fear conditioning and extinction procedures. Although the two groups did not differ on fear conditioning measures, the escitalopram group exhibited greater extinction (as measured by SCR), compared to the placebo group. If replicated, this finding would appear to support the potential role of SRIs in the prevention of PTSD. The notion that SRIs may expedite extinction learning is consistent with recent research suggesting that combining SRIs with cognitive-behavioral therapy may be beneficial (Schneier et al., 2012).

Glucocorticoids, which include stress hormones released from the adrenal cortex, have also been shown to affect extinction in rodents and symptom severity in anxiety disorders. For example, the synthetic glucocorticoid dexamethasone administered prior to or after extinction training enhanced extinction in rats (Yang, Chao, & Lu, 2006). In a randomized, double-blind, placebo-controlled study of specific phobia for heights, cortisol administered one hour before virtual-reality based exposure therapy significantly reduced fear at post-treatment and at a one-month follow-up (de Quervain et al., 2011). There is also some evidence for hydrocortisone augmentation of exposure therapy in PTSD (Surís, North, Adinoff, Powell, & Greene, 2010; Yehuda & LeDoux, 2007) and low-dose cortisol treatment of PTSD symptoms (Aerni et al., 2004).

Other agents have been shown to improve extinction measures. For example, cannabinoid agonists increased extinction in rodents (Chhatwal, Davis, Maguechak, & Ressler, 2005; Chhatwal & Ressler, 2007), and in healthy humans (Das et al., 2013; Rabinak et al., 2013). Intranasal oxytocin also appears to increase extinction in healthy humans (Acheson et al., 2013).

Another way to manipulate extinction would be to stimulate the brain regions critically involved in extinction and extinction recall. Previous research in rodents has suggested that direct electrical stimulation of the infralimbic cortex in the presence of the CS+ or fear-conditioned context reduces conditioned freezing (Milad & Quirk, 2002; Milad, Vidal-Gonzalez, & Quirk, 2004; Thompson et al. 2010). Baek, Chae, and Jeong (2012) reported similar findings with transcranial magnetic stimulation (TMS; positioned 3mm anterior to bregma) in rats. In an intriguing new study in humans, Isserles et al. (2013) administered deep TMS of the frontal cortex while subjects with PTSD were engaged in traumatic memory recall. The PTSD subjects who received TMS during traumatic memory recall showed reductions in intrusive memories and heart rate responses to the traumatic recall, whereas comparison groups did not. An earlier study (Osuch et al., 2009) reported similar findings, although the location and intensity of TMS stimulation was different from that of Isserles et al. (2013). Increasing parasympathetic nervous system output by vagus nerve stimulation increases extinction learning in rats (Peña, Engineer, & McIntyre, 2013), and may be a promising adjunct to exposure-based therapy in PTSD.

As mentioned earlier, blocking the reconsolidation of the fear memory at retrieval is a new focus of possible treatment for PTSD. Basic research suggests that reconsolidation blockade could be achieved by administering either pharmacologic agents (e.g., Diergaarde,
Schoffelmeer, & De Vries 2008; Nader et al., 2000; Pitman et al., 2011; Gamache et al., 2012) or extinction training (e.g., Monfils et al., 2009) immediately after memory reactivation (within the lability window). In line with the first method of blocking reconsolidation, Brunet et al. (2008) asked participants with PTSD to provide detailed descriptions (called “scripts”) of their traumatic events (thereby reactivating their traumatic memories) and then administered either propranolol or placebo in a randomized, double-blind design. One week later, participants’ psychophysiological responses to their own trauma scripts were measured. The group who received post-reactivation propranolol had smaller heart rate and SCR responses to the trauma scripts as compared to the placebo group. Although the study lacked a comparison group of PTSD patients who received propranolol in the absence of memory reactivation, these initial results are encouraging and warrant further examination.

Consistent with the second possible method of blocking reconsolidation, Rothbaum et al. (2012) studied the effects of administering a modified version of prolonged exposure therapy immediately (within an average of approximately 12 hours) post-trauma in an emergency department setting. Compared to an assessment-only control group, patients who received this intervention had fewer PTSD and depressive symptoms at 4 and 12 weeks post-trauma. It should be noted that previous attempts to initiate treatment immediately after trauma have not been successful (e.g., Kearns, Ressler, Zatzick & Rothbaum, 2012; Mayou, Ehlers & Hobbs, 2000), most likely due to the different type of treatment implemented (psychological debriefing vs. prolonged exposure) and/or later treatment initiation (in many cases, two or more days post trauma, which is most likely outside of the memory lability window).

7. CONCLUSION

In summary, fear conditioning and extinction has been a fruitful paradigm for understanding PTSD and other anxiety disorders. Rodent and healthy human studies have demonstrated that fear conditioning involves the amygdala and dACC creating a CS-US association. Fear extinction involves the vmPFC and hippocampus interacting to form a context-dependent CS-noUS association. Genetic studies have explained some of the variance in conditioning and extinction in healthy individuals, and may yield insights into the genetic basis of PTSD. Furthermore, fear conditioning and extinction abnormalities may be a behavioral biomarker for general vulnerability to anxiety disorders, raising the intriguing possibility that interventions that enhance extinction learning could have beneficial effects for these disorders.

Acknowledgments

The authors are supported in part by NIMH 5R01MH054636. MBV is supported by a National Defense Science and Engineering Graduate Fellowship (NDSEG).

MANUSCRIPT ABBREVIATION LIST

5-HTTLPR serotonin transporter
REFERENCES


VanElzakker et al.  Neurobiol Learn Mem. Author manuscript; available in PMC 2015 September 01.


VanElzakker, MB.; Staples, LK.; Shin, LM. The Neurocircuitry of Fear and PTSD. Vermetten, E.; Neylan, TC.; Pandi-Perumal, SR.; Kramer, M., editors. Springer-Verlag; New York: in press


Zimmerman JM, Maren S. NMDA receptor antagonism in the basolateral but not central amygdala blocks the extinction of Pavlovian fear conditioning in rats. The European Journal of Neuroscience. 2010; 31(9):1664–1670. [PubMed: 20525079]

### HIGHLIGHTS

- Pavlovian fear conditioning and extinction relate to PTSD and other anxiety disorders
- Such studies in rodents inform human psychophysiological and neuroimaging studies
- Gene and twin studies have revealed vulnerabilities to PTSD and anxiety disorders
- Treatment of PTSD and anxiety disorders use Pavlovian fear extinction as a template
FIGURE 1.
A simplified schematic of fear neurocircuitry, showing functional homology of rat and human brain regions. PL in rat and dACC in human are homologues, both project to the lateral nucleus of the amygdala. IL in rat and vmPFC in human are homologues, both project to the basal nucleus and to intercalated cells of the amygdala. The circuitry within the amygdala is shared among species. Arrowheads represent excitatory projections and circle-endings represent inhibitory projections. Within the amygdala, green shapes represent glutamatergic (excitatory) cells and red shapes represent GABAergic (inhibitory) cells. Projections from lateral nucleus to CeL, from basal nucleus to CeM, and some other connections are not shown.

FIGURE ABBREVIATIONS:
CeL = centrolateral subdivision; CeM = centromedial subdivision; dACC = dorsal anterior cingulate cortex; H = hippocampus; IL = infralimbic cortex; ITC = intercalated neurons; PL = prelimbic cortex; SC = somatosensory cortex; T = thalamus