Neural mechanisms of impaired fear inhibition in posttraumatic stress disorder

Tanja Jovanovic* and Seth Davin Norholm

1 Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, GA, USA
2 Mental Health Service Line, Veterans Affairs Medical Center, Decatur, GA USA

Posttraumatic stress disorder (PTSD) can develop in some individuals who are exposed to an event that causes extreme fear, horror, or helplessness (APA, 1994). PTSD is a complex and heterogeneous disorder, which is often co-morbid with depression, substance abuse, and anxiety disorders such as panic or social phobia. Given this complexity, progress in the field can be greatly enhanced by focusing on phenotypes that are more proximal to the neurobiology of the disorder. Such neurobiological intermediate phenotypes can provide investigative tools to increase our understanding of the roots of the disorder and develop better prevention or intervention programs. In the present paper, we argue that the inhibition of fear responses is an intermediate phenotype that is related to both the neurocircuitry associated with the disorder, and is linked to its clinical symptoms. An advantage of focusing on fear inhibition is that the neurobiology of fear has been well investigated in animal models providing the necessary groundwork in understanding alterations. Furthermore, because many paradigms can be tested across species, fear inhibition is an ideal translational tool. Here we review both the behavioral tests and measures of fear inhibition and the related neurocircuitry in neuroimaging studies with both healthy and clinical samples.

Keywords: PTSD, fear inhibition, extinction, fear neurocircuitry, neuroimaging

NEUROBIOLOGICAL INTERMEDIATE PHENOTYPES OF PTSD

Posttraumatic stress disorder (PTSD) can develop in some individuals who are exposed to an event that causes extreme fear, horror, or helplessness (APA, 1994). PTSD is considered the fourth most common psychiatric disorder, affecting 10% of all men and 18% of women (Breslau et al., 1998). The rates of lifetime PTSD are closer to 40% in high trauma populations, such as combat (Kessler et al., 1995) and low-income inner-city populations (Schwartz et al., 2005; Alim et al., 2006). Recent studies have demonstrated a steep dose–response curve between trauma frequency and PTSD symptom severity such that the more traumatic events a person experiences, the greater the PTSD symptoms (Binder et al., 2008; Mcteague et al., 2010). Even at such high prevalence rates, the relationship between trauma exposure and PTSD suggests resiliency in the majority of individuals, indicating the presence of “resilience factors” that allow trauma-related symptoms to diminish over time. These factors can be genetic, as shown by several recent gene by environment interaction studies (Binder et al., 2008; Ressler et al., 2010), or psychological, such positive social support (Charney, 2004; Norholm and Ressler, 2009).

Delineating these resilience factors is of great importance to the development of improved and personalized treatment approaches to this disorder; however, using the DSM-IV defined disorder as the phenotype under investigation raises many complications. PTSD is a heterogeneous disorder, which presents with different symptom domains, specifically, re-experiencing, avoidance and numbing, and hyper-arousal symptoms. As some patients may present higher symptoms in one domain as compared to another, a one-size-fits-all approach is often inadequate (Norrholm and Jovanovic, 2010). Furthermore, the neurobiological underpinnings of the different symptoms may not overlap (Lanius et al., 2006), suggesting that different “subtypes” of PTSD may have different treatment targets. Finally, PTSD is frequently co-morbid with other disorders, such as depression, substance abuse, and other anxiety disorders (Kessler et al., 1995). Taken together, these issues result in a complex phenotype of PTSD; one that is difficult to model in animal research and does not easily lend itself to treatment outcome studies.

Given this complexity, progress in the field can be greatly enhanced by focusing on phenotypes that are more proximal to the neurobiology of the disorder. Such neurobiological intermediate phenotypes can provide investigative tools to increase our understanding of the roots of the disorder and develop better prevention or intervention programs. Although the narrow focus cannot by definition encompass the entire spectrum of the illness, it can define targets in the neurocircuitry of the illness.

In the present paper, we argue that the inhibition of fear responses is an intermediate phenotype that is related to both the neurocircuitry associated with the disorder, and is linked to its clinical symptoms. An advantage of focusing on fear inhibition is that the neurobiology of fear has been well investigated in animal models providing the necessary groundwork in understanding alterations. Furthermore, because many paradigms can be tested across species, fear inhibition is an ideal translational tool. For example, fear-potentiated startle and inhibition of fear-potentiated startle has been tested in rodents, non-human primates, as well as humans (Myers et al., 2009). Here we review both the behavioral tests and measures of fear inhibition and the related neurocircuitry in neuroimaging studies with both healthy and clinical samples.
INHIBITION OF FEAR AS A NEUROBIOLOGICAL PHENOTYPE

Inhibition of fear responses involves learning and then appropriately responding to safety signals, i.e., the ability to discriminate between danger and safety cues and suppress fear responses in the presence of safety cues (Jovanovic et al., 2011). In the laboratory, fear inhibition can be measured by first using a fear conditioning paradigm to acquire learned fear (termed fear acquisition), which is then followed by training to inhibit learned fear responses (termed fear inhibition). Fear conditioning is based on a simple Pavlovian conditioning model in which a neutral conditioned stimulus (termed the CS; for example, a light) is paired with an aversive unconditioned stimulus (termed the US; for example, cutaneous electric shock). After a number of pairings, the association is formed so that the CS alone elicits the conditioned response (termed the CR; for example, freezing in rodents or fear-potentiated startle in humans; Pavlov, 1927). This basic model is used in animal as well as human research to investigate mechanisms of fear expression (Davis, 1990; Labar et al., 1995; Grillon and Davis, 1997; Fanselow and Ledoux, 1999; Lissek et al., 2005; Jovanovic et al., 2006). The advantage of using these paradigms is that they can be measured with peripheral outcomes such as the skin conductance or startle responses, which are non-invasive but offer physiological measures of fear conditioning. In this review we will describe fear inhibition as the reduction of fear responses in the presence of safety cues which is a manifestation of the underlying inhibitory neurocircuitry.

There are two laboratory models that have been primarily used for behavioral testing of fear inhibition in animals and humans: extinction and differential conditioning. Whereas fear acquisition refers to learning that something is dangerous, extinction is a mechanism by which an individual learns that something that was previously dangerous has become safe. In fear extinction paradigms, a stimulus that was previously paired with an aversive stimulus (the CS+) is then repeatedly presented without the US, so that it no longer elicits a fear response [cf. (Myers et al., 2006; Norrholm et al., 2006), see Figure 1A]. In a basic differential conditioning paradigm, the above CS+ pairing is intermingled, at the time of training, with a separate stimulus (CS−). The CS− does not co-occur with an aversive stimulus, and thus represents safety, or inhibition of fear. This paradigm involves a simple discrimination between the danger and safety cues (see Figure 1B), and is the one most commonly used in human fear conditioning research (Lissek et al., 2005). More complex tasks, such as conditional discrimination, Figure 1C (in which there is an element of the conditioning stimulus that is shared between the CS+ and CS−), and stimulus generalization, Figure 1D (in which there is a perceptual gradient of stimuli between the CS+ and CS−), are designed to capture more subtle variation in fear-inhibition processes.

Conditioned inhibition involves a variation of discrimination in which the danger cue (i.e., CS+) is not reinforced when preceded by (or combined with) a second cue, usually termed X, so that CS+ alone, CS/ X−). Although the X cue should be conditioned to designate safety, it is vulnerable to second-order conditioning effects and limited by configural processing (Myers and Davis, 2004). This paradigm has been used in several animal studies (Falls et al., 1997; Gewirtz et al., 1997), but rarely in human studies (Grillon and Ameli, 2001) due to the above issues. Conditional discrimination, a modification

<table>
<thead>
<tr>
<th>A</th>
<th>EXTINCTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase I:</td>
<td>Phase II:</td>
</tr>
<tr>
<td><img src="https://www.frontiersin.org" alt="Image" /></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B</th>
<th>SIMPLE DISCRIMINATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase I:</td>
<td>Phase II:</td>
</tr>
<tr>
<td><img src="https://www.frontiersin.org" alt="Image" /></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C</th>
<th>CONDITIONAL DISCRIMINATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase I:</td>
<td>Phase II:</td>
</tr>
<tr>
<td><img src="https://www.frontiersin.org" alt="Image" /></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D</th>
<th>STIMULUS GENERALIZATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase I:</td>
<td>Phase II:</td>
</tr>
<tr>
<td><img src="https://www.frontiersin.org" alt="Image" /></td>
<td></td>
</tr>
</tbody>
</table>

**FIGURE 1** Schematic depictions of four different fear-inhibition paradigms; (A) Extinction, (B) Simple Discrimination, (C) Conditional Discrimination, and (D) Stimulus Generalization. CS+ = Reinforced Conditioned Stimulus; CS− = Non-reinforced Conditioned Stimulus. In fear extinction paradigms (A), a stimulus that was paired with an aversive stimulus (the CS+) during the initial phase is then repeatedly presented without the US in the next phase of the experiment, so that it no longer elicits a fear response. In the simple discrimination paradigm (B), the CS+ is intermingled with a separate stimulus that is not paired with the US (CS−). In a conditional discrimination experiment (C), the US occurs depending on the presence of the CS+ when it is combined with a neutral (X) cue, which is also combined with the CS− to predict the absence of the US. In the next phase of the experiment, presentation of CS+ and CS− together, without the shared cue (X) results in a reduced fear response compared to the response to CS+. In a stimulus generalization experiment (D), after the initial conditioning, the second phase of the experiment presents iterations of CSs that differ from the CS+ in small increments. The number of degrees required for discrimination is an indication of cue overgeneralization.
of the conditioned inhibition design, avoids some of these limitations (Myers and Davis, 2004). In a conditional discrimination experiment, an aversive event occurs depending on the presence of the CS+ when it is combined with a neutral (X) cue (Wagner and Rescorla, 1972). This cue is also combined with the CS− to predict the absence of the US (Figure 1C). In a critical subsequent test trial, presentation of CS+ and CS− together, without the shared cue (X) results in a reduced fear response compared to the response to CS+. This is referred to as the transfer test, when the inhibitory properties of the CS− are actively transferred to the combined CS+/CS− cue (Myers and Davis, 2004; Jovanovic et al., 2005).

The current review will focus on extinction, simple discrimination, and conditional discrimination methods of fear inhibition, as these have been more thoroughly investigated using psychophysiological techniques. The final paradigm, stimulus generalization is a very novel approach to studying fear inhibition and has only been recently used in patients with panic disorder (Lissek et al., 2010), and there are no published studies to date that have tested stimulus generalization in PTSD. This paradigm is promising because it assesses both the subjects’ ability to detect subtle differences between danger and safety on a continuum, and their ability to show reduction of fear once the discrimination occurs (Lissek et al., 2010).

Both extinction tests and differential conditioning paradigms focus on active suppression of fear responses through learned safety signals; while fear itself may only involve subcortical areas of the brain located primarily in the limbic circuitry, safety signals may require a cognitive, cortical component (Bremner et al., 2005; Weike et al., 2008). This premise is supported by data from our lab showing that awareness of the association between the CS and the US is necessary for inhibiting fear responses (Jovanovic et al., 2006). Furthermore, a recent study by Weike et al. (2008) examined the temporal domain of fear conditioning with a danger and safety signal and found that safety signal processing was slower than danger processing. The authors argued that top-down cognitive processes are involved in responses to safety signals which accounts for the latency in response.

We have recently used extinction, simple discrimination, and conditional discrimination paradigms in a highly traumatized civilian population from inner-city Atlanta (Jovanovic et al., 2010a, b; Norrholm et al., 2011). Data from our study on extinction (Norrholm et al., 2011) suggest that the early phase of extinction is predicted by the level of fear expression to the CS+ (i.e., the danger signal) at the end of acquisition. It is this fear expression during early extinction that is exaggerated in PTSD subjects compared to traumatized non-PTSD controls (see Figure 2A). On the other hand, a high degree of fear remaining during late extinction is related to impaired inhibition, as it is best predicted by responses to the CS− (i.e., safety signal) at the end of acquisition (Norrholm et al., 2011). Figure 2B shows simple discrimination between the CS+ and the CS− during late acquisition between PTSD subjects and controls. Although PTSD subjects are slower in developing the discrimination, by the final phase of conditioning both groups show higher levels of fear-potentiated startle to the CS+ than the CS−; however, PTSD subjects demonstrate higher levels of fear to both stimuli (Norrholm et al., 2011). The final paradigm, conditional discrimination, measures fear inhibition to the safety signal by testing the reduction of fear-potentiated startle on the transfer test. We used this paradigm in combat veterans (Jovanovic et al., 2009) and traumatized civilian population (Jovanovic et al., 2010b) and have found that individuals who met criteria for PTSD had higher potentiation of the startle response to the CS− than traumatized controls and did not transfer safety on the test trial (Figure 2C).

While some data with combat veterans suggest that impaired fear inhibition may be an acquired trait (Milad et al., 2008), that is associated with current symptom severity (Jovanovic et al., 2009), other studies have reported that heightened fear responses and decreased inhibition of fear may be predictors of the disorder. A prospective study of police academy cadets found that greater skin conductance responses to threatening stimuli and slower habituation prior to trauma exposure were predictive of PTSD symptom severity after trauma exposure (Pole et al., 2009). A similar prospective study with firefighters found that reduced extinction of fear conditioned responses examined before the index trauma explained almost a third of the symptoms in later traumatized
disorder and contributes to the maintenance of the disorder, while decreased ability to inhibit fear is a risk factor for developing the disorder. It is possible that a recent study of Vietnam veterans and their twins found that PTSD subjects did not show the same impairment. It is possible that a decreased ability to inhibit fear is a risk factor for developing the disorder and contributes to the maintenance of the disorder, while decreased extinction retention on the day after acquisition and extinction occurred at lower magnitudes in PTSD subjects than in controls (Milad et al., 2008). Furthermore, impaired retention of extinction appeared to be an acquired trait related to the disorder since the twins of the PTSD subjects did not show the same impairment. It is possible that a decreased ability to inhibit fear is a risk factor for developing the disorder and contributes to the maintenance of the disorder, while decreased extinction retention on the day after acquisition and extinction occurred at lower magnitudes in PTSD subjects than in controls (Milad et al., 2008).

Extinction learning requires inhibition of the fear circuitry by the prefrontal cortex (PFC; Phelps et al., 2004); whereas discrimination between danger and safety cues, and recall of extinction may also require activation of the hippocampus (Milad et al., 2007b). Given that both extinction and differential conditioning are dependent on the PFC, this review will focus on this region as a primary locus of fear-inhibition neurocircuitry.

INHIBITION NEUROCIRCUITRY AS A TARGET OF INVESTIGATION

The PFC has long been thought to play a role in behavioral inhibition. More than a decade ago, animal studies reported that lesions of the medial PFC (mPFC) prior to original fear conditioning retard extinction to a tone (Morgan et al., 1993). Recent studies have demonstrated that neurons in the PFC may have inhibitory action on the amygdala (Grace and Rosenkranz, 2002; Phelps et al., 2004). The PFC can be subdivided into medial and orbitofrontal PFC. The anterior cingulate cortex (ACC), which is also part of the PFC, has both rostral and dorsal components which may play different roles in the expression and inhibition of fear, as will be discussed in greater detail below. Figure 3 shows the medial regions of the PFC most involved in inhibitory processes (Figure 3A), including the ventral and dorsal PFC and the ACC (Figure 3B).

A well known and frequently used example of a more complex inhibition task is the Stroop effect task, where the meaning of a word (such as the word “red”) is in conflict with the color in which it is shown (for example, in blue ink). In this task, the subject is instructed to state the color of the ink while ignoring the interference from the word. Due to the conflict between the color and the word, reaction times are delayed, providing a measure of the cognitive inhibition (Stroop, 1935), and activating the ACC (Pardo et al., 1990; Bremner et al., 2004). The Stroop task can also be adapted to use with emotion-relevant stimuli, in which the emotional content of a word competes with the cognitive content and must be ignored. This task also activates the ACC, but in an area distinct from the strictly cognitive interference tasks (Whalen et al., 1998, 2006). Emotionally relevant stimuli appear to be processed by the rostral or subgenual area of the ACC (Shin et al., 2005), which is anterior to the genu of the corpus callosum. Furthermore, this specific region of the PFC, the rostral ACC, is involved in amygdala regulation (Etkin et al., 2006).

Neuroimaging studies using fear conditioning paradigms demonstrate that fear acquisition and extinction of fear also activate the PFC, specifically the ventromedial (vmPFC; Phelps et al., 2004; Reinhardt et al., 2010). Recent developments in the spatial resolution of neuroimaging techniques have resulted in more fine-tuned examinations of this area of the brain. As mentioned above, the rostral or subgenual regions of the ACC are activated during the presentation of emotional stimuli; these areas are also activated during the regulation of fear (Phelps et al., 2004; Schiller et al., 2008). There are several lines of evidence that this region of the vmPFC is associated with inhibition of fear. For example, fMRI data
is dysregulated in PTSD patients. Evidence suggests that a hallmark in the expression of negative affect and not simply activated by this area has been associated with fear as well as other noxious stimuli acquisition, rather than by the fear itself. However, given that this area is also implicated in cognitive tasks (Shin et al., 2007), it may be activated by the active learning that occurs during fear acquisition, rather than by the fear itself. However, given that this area has been associated with fear as well as other noxious stimuli such as pain (Vogt et al., 2003), it may be more centrally involved in the expression of negative affect and not simply activated by general learning.

Several studies have indicated that this inhibitory neurocircuit is dysregulated in PTSD patients. Evidence suggests that a hallmark of PTSD neurobiology is exaggerated amygdala activity during fearful stimulation coupled with reduced top-down control of the amygdala by the PFC (Liberzon et al., 1999; Rauch et al., 2000, 2006; Shin et al., 2004; Liberzon and Martinis, 2006). Furthermore, functional neuroimaging studies that have examined connectivity between PFC and the amygdala have demonstrated impaired inhibition of the amygdala in PTSD (Lanius et al., 2004). The emergence of MRIs with greater special resolution allows for more precise descriptions of these neural substrates. A recent meta-analyses of imaging studies during emotion processing in PTSD, social anxiety, and specific phobia indicated that the rostral ACC is less active in PTSD patients relative to controls; an area may be a risk factor for psychopathology: a recent study of combat veterans (Shin et al., 2001). Decreased activation of this area may be due to treatment efficacy, but rather to a practice effect from the difficulty of testing learning in a repeated design. For example, improvements in fear inhibition after treatment may not be due to treatment efficacy, but rather to a practice effect from patients remembering the previously administered training paradigms. A recent study with fear acquisition and fear extinction tests spaced 12 weeks apart demonstrated good test–retest reliability on patients remembering the previously administered training paradigms (Jovanovic et al., 2009). In addition, impaired fear inhibition that manifests itself as disrupted extinction recall may represent acquired traits of the disorder (Milad et al., 2008). The difference between these paradigms may lie in the involvement of the hippocampus in the latter (Milad et al., 2007c); which early studies have found to be decreased in PTSD subjects (Bremner et al., 1995) possibly as a result of the trauma and associated stress (Bremner, 2001). However, some have argued that smaller hippocampal volume is a risk factor for PTSD (Pitman, 2001; Gilbertson et al., 2002). These issues can only be resolved with a prospective study examining neural volume before and after onset of PTSD. Regardless of whether fear inhibition is impaired prior to PTSD or develops as part of the disorder itself, it may be modifiable with treatment.

Clinically, fear inhibition has not yet been tested with regard to treatment response. The lack of these studies is partly due to the difficulty of testing de novo learning in a repeated design. For example, improvements in fear inhibition after treatment may not be due to treatment efficacy, but rather to a practice effect from patients remembering the previously administered training paradigms. A recent study with fear acquisition and fear extinction tests spaced 12 weeks apart demonstrated good test–retest reliability on these measures (Zeidan et al., 2011); this is a crucial first step to developing treatment outcome measures.

Although treatment outcome and psychophysiological fear-inhibition measures have not been actively examined concurrently, several studies have begun to investigate the relationship between pre-treatment neural function and subsequent treatment response. In PTSD, one study revealed that larger rostral ACC
volume predicted positive outcomes to cognitive-behavior therapy (Bryant et al., 2008). Furthermore, this brain area has been linked to treatment response in depression (Pizzagalli et al., 2001; Pizzagalli, 2011). The ventral area immediately below the corpus callosum has been used as a target in deep brain stimulation to relieve depression (Holtzheimer and Mayberg, 2010), with long-term positive outcomes (Kennedy et al., 2011).

There is an emerging body of literature assessing structural and functional changes in the neural underpinnings of PTSD with treatment. An early study using single photon emission computed tomography (SPECT) imaging pre- and post-treatment with selective serotonin reuptake inhibitors (SSRIs), found significant changes in ACC and hippocampus after 12 weeks of treatment (Carey et al., 2004). More recent studies using fMRI before and after psychotherapy for depression have indicated changes in prefrontal areas (Dichter et al., 2010; Ritchey et al., 2011).

**FUTURE DIRECTIONS: CAN WE INCREASE NEURAL INHIBITION OF FEAR?**

Several exciting and novel avenues have been revealed for the further exploration and development of neurobiologically based, translational studies of PTSD and trauma-related disorders. We are currently well-poised to investigate these avenues as a means of developing better diagnostic tools based on novel neurobiological intermediate phenotypes. An obvious first step is to test fear inhibition before and after treatment in order to improve existing predictors of treatment response. This will also allow for the exploration of potential individual differences that contribute to positive treatment outcomes and aid in personalization on treatment strategies. Another goal is to track treatment efficacy in responders.

In addition to increased exploration of the putative fear-inhibition phenotype with respect to treatment, we also need to investigate treatment-related changes in brain neurocircuitry and structure. As previously described, fMRI has been successfully employed in depressed patients with significant treatment results (Ritchey et al., 2011). Given the specificity of the brain areas that are related to fear inhibition in PTSD, namely, the rostral ACC, simple inhibition tasks such as the Go/NoGo can be administered before and after treatment to detect changes in the volume, activity, and connectivity of this area. Furthermore, the use of fMRI tasks in concert with fear-inhibition paradigms (e.g., extinction and differential conditioning) with PTSD patients is in its infancy and we anticipate several emerging studies using these methods.

Finally, future studies should be focused on strengthening the activation of the rostral ACC, especially during the presentation of fear-related stimuli. The aforementioned fMRI studies from the depression literature suggest that the PFC represents a relatively neuroplastic brain region with the potential for treatment-related modifications in activity (e.g., successfully attenuating amygdala-driven fear responses); these findings offer great promise to improving on the available treatments for PTSD. Novel treatment approaches have employed computerized tasks in order to train patients to generate the anticipated (or therapeutic) response. For example, attention bias training has been used successfully with many anxiety disorders to train patients’ attention away from negative stimuli and, as such, significantly decrease symptom severity (Reese et al., 2010; Rozenman et al., 2011). Similar computerized tasks or games can be devised to bolster fear-inhibition training. For example, exposure therapy, which is one of the most successful psychotherapeutic approaches to PTSD, may in fact produce its effects by increasing fear inhibition. Given that exposure therapy is based on extinction learning, which activates the vmPFC, this premise would provide a likely neural mechanism of action. Facilitating extinction, either pharmacologically or behaviorally, may produce therapeutic, beneficial modifications to underlying neural connectivity and thus increase inhibition of fear.

**ACKNOWLEDGMENTS**

Dr. Jovanovic has funding from NIMH (F32 MH070129). Dr. Norrholm has research support from the National Alliance for Research on Schizophrenia and Depression (NARSAD), the Department of Defense (DOD)/Congressionally Directed Medical Research Program (CDMRP, award #W81XWH-08-2-0170), and the Emory University Research Committee.

**REFERENCES**


Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 27 April 2011; paper pending published: 27 June 2011; accepted: 11 July 2011; published online: 25 July 2011.


Copyright © 2011 Jovanovic and Norrholm. This is an open-access article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.