

# Inhibition of Fear by Learned Safety Signals: A Mini-Symposium Review

John P. Christianson,<sup>1</sup> Anushka B. P. Fernando,<sup>2</sup> Andy M. Kazama,<sup>3</sup> Tanja Jovanovic,<sup>4</sup> Linnaea E. Ostroff,<sup>5</sup>  
and Susan Sangha<sup>6</sup>

<sup>1</sup>Department of Psychology and Neuroscience, University of Colorado, Boulder, Colorado 80309-0345, <sup>2</sup>Department of Psychology and Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge CB2 3EB, United Kingdom, <sup>3</sup>Department of Developmental Cognitive Neuroscience, Yerkes National Primate Research Center, Emory University, Atlanta, Georgia 30322, <sup>4</sup>Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, Georgia 30349, <sup>5</sup>Center for Neural Science, New York University, New York, New York 10003, and <sup>6</sup>Ernest Gallo Clinic and Research Center, University of California San Francisco, Emeryville, California 94608

Safety signals are learned cues that predict the nonoccurrence of an aversive event. As such, safety signals are potent inhibitors of fear and stress responses. Investigations of safety signal learning have increased over the last few years due in part to the finding that traumatized persons are unable to use safety cues to inhibit fear, making it a clinically relevant phenotype. The goal of this review is to present recent advances relating to the neural and behavioral mechanisms of safety learning, and expression in rodents, nonhuman primates, and humans.

## Introduction

Despite numerous advances, the pathophysiology of post-traumatic stress disorder (PTSD) is not sufficiently understood and current treatments are not always therapeutic. A hallmark feature of PTSD is the heightened expression of fear or anxiety in environments where it is not appropriate. This symptom has been conceptualized as a generalization of the fear conditioned during the traumatic experience that becomes resistant to extinction (Rauch et al., 2006). While this model has significant empirical support, emerging evidence suggests that in addition to extinction learning, safety learning is impaired (Jovanovic et al., 2012a). As opposed to danger learning where a cue is paired with aversive stimulation, safety learning involves associating distinct environmental stimuli (safety signals) with the nonoccurrence of aversive events. This is distinguishable from the phenomenon of fear extinction. During extinction, the danger cue is presented without aversive reinforcement and over time a new association forms that the cue predicts the absence of the aversive event. Thus, the extinction learning introduces a new memory that competes with the original danger association, which results in the inhibition of fear (Bouton, 2004). In contrast, learned safety signals inhibit fear responses to cues that are normally still paired

with an aversive event when the safety signal is not present. As such, safety signals are only learned when the subject expects danger but it does not occur. More fundamental to the clinical importance of safety learning, the distinction between safe and dangerous circumstances is critical to survival in all animals, and it is this process that is impaired in PTSD. Thus, identifying the mechanisms of safety learning represents a significant goal for basic neuroscience that should inform future prevention and treatment of PTSD and other anxiety disorders.

A “safety signal” is a specific class of conditioned inhibitor; as a result of pavlovian conditioning, it prevents or reduces the expression of fearful behaviors normally observed in the presence of an excitatory conditioned stimulus (CS) that had been paired with aversive unconditioned stimuli (USs). Thus, the first requirement for a safety signal is that it must come to inhibit the conditioned fear response as a result of learning—stimuli that interfere with conditioned responses without training are called external inhibitors. In fact, many of the stimuli commonly used in fear-conditioning experiments may act as external inhibitors to fear responses such as startle and freezing (Myers and Davis 2004). For example, in one of our laboratories white noise was found to reduce behavioral freezing to a danger CS without prior conditioning (J. Christianson, unpublished observations). A stimulus intended for inhibitory conditioning should be excluded from a safety-learning protocol if evidence for external inhibition is found.

Safety signals have been trained in numerous procedures (Fig. 1). The critical procedure for observing fear inhibition by a safety signal is called a “summation test” and was first demonstrated by Hammond (1967). The example in Figure 1 is the result of A+/B– training in which A trials concluded with a mild foot-shock and B trials did not. The predicted result is that cue A becomes a danger CS and cue B becomes a safety signal. Summa-

Received July 9, 2012; revised Aug. 17, 2012; accepted Aug. 29, 2012.

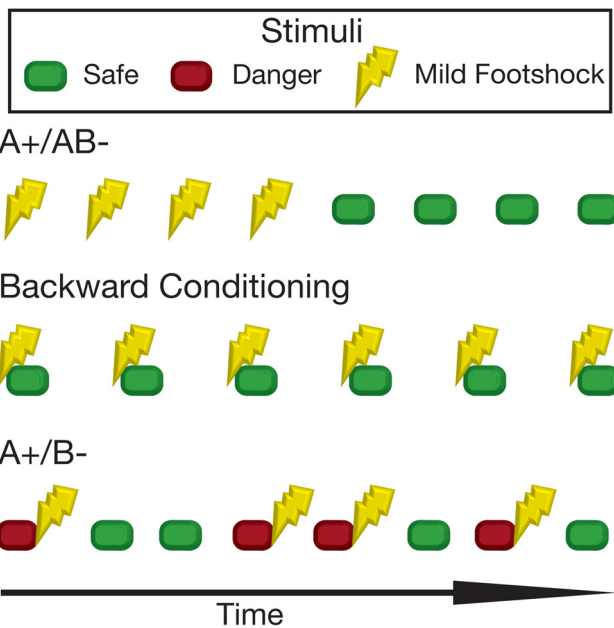
This work was supported by National Institutes of Health Grants MH093412, MH070129, MH092576, MH47840, MH088985, MH58846, MH086947, and MH083583; the National Alliance for Research on Schizophrenia and Depression; Medical Research Council Case Studentship; the Wellcome Trust; the State of California for Medical Research on Alcohol and Substance Abuse through the University of California at San Francisco to Dr. Patricia H. Janak; Yerkes Base Grant RR-00165; and National Center for Research Resources Grant P51RR165, currently supported by the Office of Research Infrastructure Programs/Grant OD P51OD11132.

Correspondence should be addressed to John P. Christianson, Department of Psychology and Neuroscience, University of Colorado, Boulder, CO 80309-0345. E-mail: john.christianson@colorado.edu.

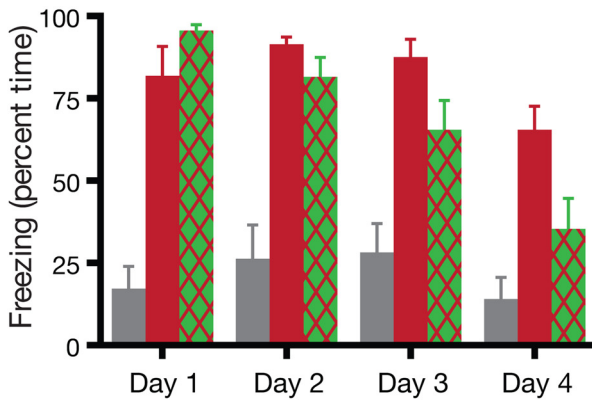
DOI:10.1523/JNEUROSCI.3340-12.2012

Copyright © 2012 the authors 0270-6474/12/3214118-07\$15.00/0

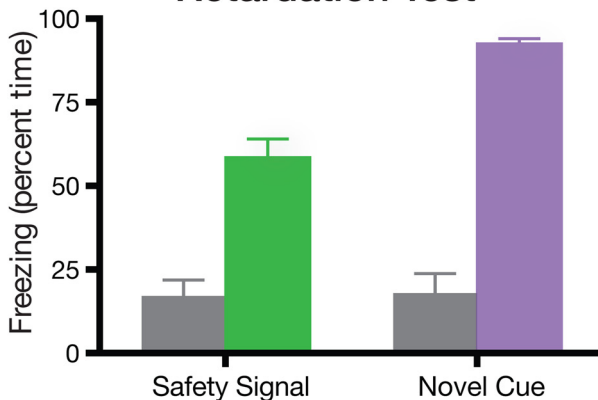
### Safety Learning Procedures



### Summation Tests



### Retardation Test



**Figure 1.** Examples of safety-learning procedures and tests. Top, Among the many procedures used to condition a safety signal, most are related to the A+/AB-, backward-conditioned, or A+/B- (discrimination training) designs. A+/AB- conditioning (Rogan et al., 2005; Ostroff et al., 2010) involves separating the aversive cue from the putative safety signal in time by either clustering footshocks together (as shown) or interspersing safety signals between shocks with long intershock intervals. In a common experiment, the A cue is the

tion tests involve presentations of the putative safety signal in compound with a conditioned excitor. After A+/B- training, a significant reduction in behavioral freezing during AB is seen. Once learned, it is difficult to retrain a safety signal as a danger CS. Evidence for this phenomenon occurs after the safety signal is paired with the aversive stimulus (Rescorla, 1969b). Upon later presentation of the safety signal, less conditioned fear is observed compared with that equally conditioned to a novel CS. Rescorla (1969b) described this as a “retardation” test because the observation is that new learning to the conditioned inhibitor appears to be delayed.

Pavlov (1927) and Konorski (1948) made early and significant experimental and theoretical contributions to conditioned inhibition. Later, a paper by Rescorla (1969b) identified the boundary conditions for establishing a conditioned inhibitor that have shaped modern behavioral and neuroscience research. The authors point interested readers to the aforementioned seminal works and a volume by Miller and Spear (1985) that discuss the nuances of conditioned inhibition. This review aims to introduce a broad audience to recent advances concerning the conditions under which safety signals are learned and used in rodent, non-human primate, and human subjects. Here we highlight the novel contributions of the investigators who participated in a Mini-Symposium panel titled “Inhibition of Fear by Learned Safety Signals” at the Annual Meeting of the Society for Neuroscience, as well as the noteworthy discoveries of others.

### Neural circuitry of safety signals

Considering that information about safety is only relevant when one anticipates danger, knowledge about the circuitry of danger is a prerequisite to understanding safety. Fearful behaviors such as enhanced startle, behavioral freezing, social avoidance, and autonomic arousal are the product of a well understood neural circuit. When presented with a US, such as a footshock, thalamic neurons transmit sensory information regarding the US and co-occurring environmental stimuli to the lateral amygdala (LA) where plasticity occurs, thereby linking the US with environmental stimuli that now predict danger. A series of excitatory relays from the lateral to the basal amygdala (BA) and then to the central nucleus of the amygdala (CeA) mediate the expression of fearful behaviors. This is an oversimplified schema (for excellent reviews, see Kim and Jung, 2006; Tronson et al., 2012), but it nevertheless captures a wealth of data. The majority of research focused on neural mechanisms of safety learning has centered on

conditioning context and the B cue is a discrete stimulus (light or sound). In backward conditioning, the safety signal occurs just after termination of shock and so predicts the intershock period (Christianson et al., 2008). In A+/B- conditioning, trials are presented in which a discrete cue A is always followed by shock and a discrete cue B is not. In A+/B-, the context provides some expectation of shock during the B trials. Often a third cue (X) is added to both types of trials to transfer some expectation of fear to the B trials (Myers and Davis, 2004). Middle, Rats were trained on A+/B- (15 trials of each type) daily for 4 d. Twenty-four hours after training, rats returned to the training context and behavioral freezing was observed over 3 min. The first minute served as a baseline (gray bars); in the second minute the A cue was presented (red bars), and in the third minute the B cue was superimposed upon A (green and red hashed bars). Over some days, the B cue becomes a safety signal and inhibits freezing in the presence of A. Bottom, After 4 d of A+/B- training, rats were then given 2 shock pairings of either the safety signal (B) or a novel cue. In a fear recall test 24 h later (baseline freezing in gray bars), greater fear was observed to the novel cue (purple bar) than to the safety signal (green bar), indicating that safety training slowed new fear acquisition (example results are from J.P. Christianson, unpublished data). These data are intended to provide an instructive example for readers unfamiliar with safety signal procedures, and our laboratories have used all of these approaches to identify the neural mechanisms of safety learning.

a simple hypothesis: safety signals must inhibit the output of the amygdala (for exceptions, see Weirtelak et al., 1992; Watkins et al., 1998). Amygdala neurons in the fear circuit are under the inhibitory control of local GABAergic interneurons (Ehrlich et al., 2009), the medial intercalated neurons (Amano et al., 2010), and the infralimbic ventromedial prefrontal cortex (vmPFC) (Milad and Quirk, 2002). In addition to these known inhibitory circuits, the amygdala receives input from most cortical regions, hippocampus, striatum, and brainstem nuclei (Sah et al., 2003); thus, several putative pathways may contribute to safety signal processing.

Modern studies of safety learning using lesion and pharmacological manipulations offer a litany of brain structures deemed insufficient to affect safety learning. A significant effort made by the laboratory of Michael Davis and his students led to reports of null effects in nucleus accumbens (Josselyn et al., 2005), CeA (Falls and Davis, 1995), vmPFC (Gewirtz et al., 1997; Christianson et al., 2008), hippocampus (Heldt et al., 2002), and perirhinal cortex (Falls et al., 1997). None of these manipulations were sufficient to impair safety signaling. However, lesions to the auditory thalamus and superior colliculus, sensory inputs to the LA, did disrupt conditioned inhibition of fear (Waddell et al., 2003; Heldt and Falls, 2006).

The LA receives direct inputs from auditory areas that process stimuli commonly used to train fear and safety associations. As the animal learns to associate new meanings with these stimuli, changes in these inputs can be studied. Rogan et al. (2005) recorded synaptic responses to an auditory tone in the LA of awake, behaving mice. As had been seen previously, tone-evoked responses were enhanced when animals were fear conditioned to the tone (Rogan et al., 1997). Interestingly, these responses weakened when the tone was trained as a safety signal. These data indicate that a safety signal may not only inhibit the output of the amygdala complex, but also reduce a sensory cue's ability to excite the LA. This could explain the delay in acquiring a fear association to a previously established safety signal, as the relevant inputs must be increased from below baseline. Another comparison of LA synapses after fear or safety training found that synapse size increased with fear conditioning and decreased with safety conditioning (Ostroff et al., 2010). Interestingly, while there was additional evidence of changes in synapse morphology and density with fear conditioning (Ostroff et al., 2012), only reduced synapse size was seen with safety conditioning.

Simply reducing the potential of a stimulus to excite fear responses is not sufficient to establish a safety signal. It is therefore unlikely that weakening of LA synapses is the central mechanism of safety signal learning (although strengthening of LA synapses is undoubtedly a central mechanism of fear conditioning). In addition to recording in the LA, Rogan et al. (2005) also recorded tone-evoked synaptic responses in the striatum. Here, responses were enhanced with safety conditioning and weakened with fear conditioning. This indicates that the safety signal excites a region associated with approach and reward, providing it with an activating function that may be involved in switching behavior from defensive avoidance to approach when the safety signal is provided.

Safety signals possess rewarding qualities (discussed in detail below), and this led Pollak et al. (2008) to test whether safety training would act as a behavioral "antidepressant." Indeed, safety training led to changes in forced swimming behaviors and anhedonia that occur with chronic fluoxetine treatment. As with chronic fluoxetine, safety training increased neurogenesis in the dentate gyrus of the hippocampus and x-irradiation, a treatment

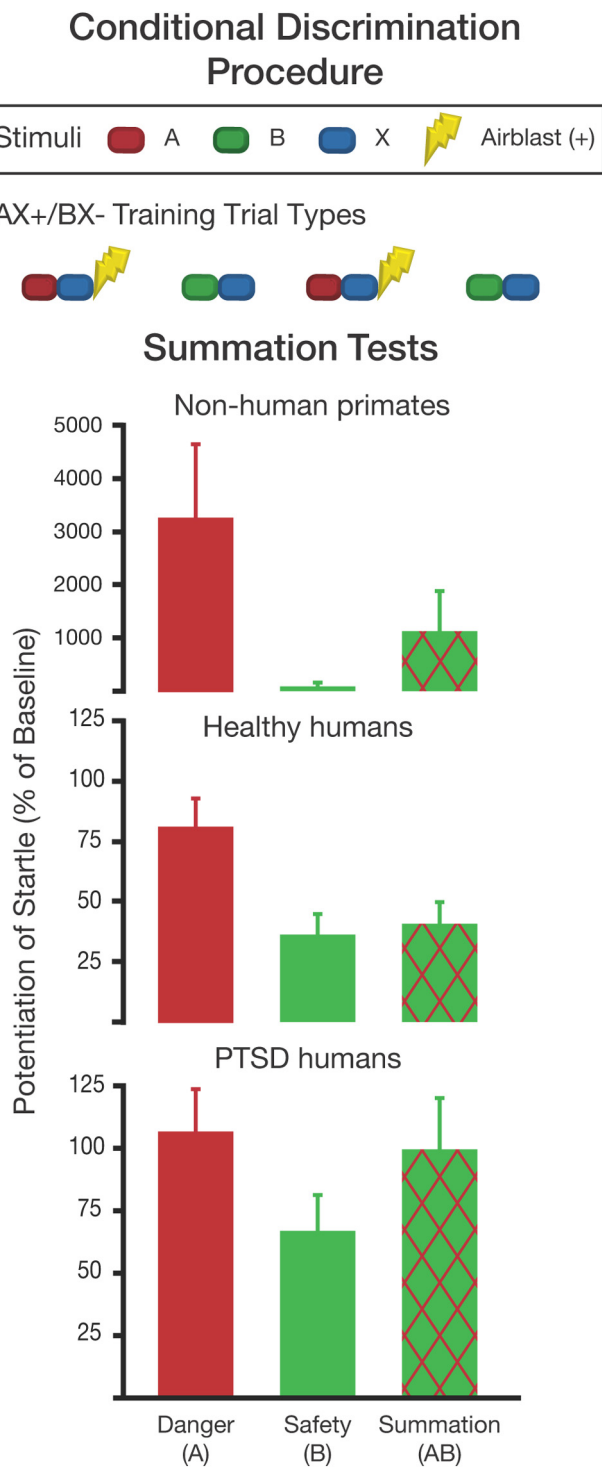
that prevents neurogenesis, prevented the antidepressant-like effects of safety training. Important to the discussion of safety signal mechanisms, the data of Pollak et al. (2008) imply that the hippocampus contributes to safety signal processing; however, additional studies will be required to determine whether its role is in learning or in regulation of emotion.

As safety signals are potent inhibitors of fear, there exists a hypothesis that safety signals can mitigate the consequences of intermittent stressors (Weiss, 1971; Mineka et al., 1984). In fact, providing a safety signal during an intermittent shock stressor significantly attenuated (1) the behavioral freezing response during the stressor; (2) the induction of Fos protein in the LA, posterior BA, and bed nucleus of the stria terminalis (Christianson et al., 2011); and (3) the long-term anxiogenic effect of the stressor (Christianson et al., 2008). Since lesions to the amygdala completely abolish the consequences of intermittent shock stress (Maier et al., 1993), cortical structures projecting to the amygdala nuclei were considered candidates for these safety signal effects. The posterior insular cortex (IC) receives multimodal somatosensory input (Nieuwenhuys, 2012), exhibits convergent responses to simultaneous multisensory stimulation (Rodgers et al., 2008), and projects to the amygdala (McDonald et al., 1999). These characteristics led Christianson et al. (2008, 2011) to hypothesize that the IC would be involved in learned safety. Indeed, both excitotoxic lesions (Christianson et al., 2008) and reversible pharmacological lesions (Christianson et al., 2011) completely abolished the stress-mitigating effect of the safety signal. This discovery added an important new candidate structure involved in emotional regulation. Importantly, the IC is a site of functional and structural abnormalities in anxiety and PTSD (Paulus and Stein, 2006; Hughes and Shin, 2011).

### Safety signal learning in human and nonhuman primates

An advantage of focusing on fear inhibition by safety signals is that the known neurobiology of fear provides the necessary groundwork to understand fear-related mental disorders. Excessive fear and anxiety, along with an inability to overcome these emotions, are some of the defining characteristics of many anxiety disorders, such as phobias, panic disorder, and PTSD. Overgeneralization of trauma-related stimuli or situations (i.e., an impaired ability to discriminate between danger and safety cues (Jovanovic et al., 2009, 2010) can lead to hypervigilance and exaggerated physiological responses that are part of the PTSD clinical presentation. For example, combat veterans with PTSD may not be able to suppress fear in response to a previously learned fearful cue (e.g., helicopter sound), even when surrounded by many safety signals (e.g., a different time and place from the combat zone).

Laboratory paradigms that specifically test safety signal processing offer an objective assessment of the clinical impairment that does not rely solely on the patient's self-report of symptoms. Jovanovic et al. (2005) have translated a conditional discrimination procedure (AX+/BX−) based on a rodent model (Fig. 2) in which healthy individuals readily acquire excitatory associations with cue A when paired with X, whereas B paired with X becomes a safety signal. Safety signal learning is tested by pairing A with B on summation test trials, which show a decreased fear response compared with AX (Fig. 2). Jovanovic et al. (2009, 2010) have used this paradigm in PTSD patients from combat (Jovanovic et al., 2009) and civilian populations (Jovanovic et al., 2010), and both demonstrate impaired fear inhibition; Figure 2 shows data from both PTSD samples combined. Importantly, this phenomenon is unique to trauma-related disorders and is not seen in



**Figure 2.** Conditional discrimination procedure with human and nonhuman primates. Top, The AX+/BX- conditioning phase includes presentations of A and X together predicting the US (airblast is used in the primate studies), while B and X together predict the nonoccurrence of the US (Myers and Davis, 2004). Bottom, A and B are presented separately and then together during the post-training summation test; B acts as an inhibitor to reduce fear in the AB compound compared with AX. Intact rhesus macaques were trained on AX+/BX- and show reduction in fear on the AB compound (Kazama et al., 2012). Similarly, healthy humans show a significant reduction of fear to AB relative to A. PTSD subjects discriminated between A and B, but they did not inhibit startle during AB trials (data were combined from Jovanovic et al., 2010, 2012b).

comorbid mental illnesses such as depression (Jovanovic et al., 2010). In PTSD patients, safety signal deficiencies may appear as early as 30 d and as late as 10+ years after trauma exposure, indicating that it is a persistent biomarker of psychopathology (Jovanovic et al., 2012b). The underlying neurobiological mechanisms for impaired safety signal processing in PTSD may involve an interaction between hyperactivity of the amygdala and impaired top-down emotional control by the vmPFC (Rauch et al., 2006). Hypoactivity and structural differences in the vmPFC are consistently observed in PTSD populations (Corbo et al., 2005; Etkin and Wager, 2007; Hughes and Shin, 2011). Thus, the vmPFC is hypothesized to be a site of pathophysiology in PTSD and may contribute to impairments in safety signal use.

As noted above, lesions to vmPFC in rat do not interfere with safety signals (Gewirtz et al., 1997; Christianson et al., 2008), weakening the vmPFC hypothesis. However, translation to the clinical model is problematic due to structural differences between the rodent and human frontal cortex (for comparison, see Ongür and Price, 2000). In contrast to rodents, rhesus macaques share similar frontal cortices to humans, making them indispensable translational models. In the nonhuman primate version of the AX+/BX- fear-potentiated startle paradigm (Winslow et al., 2008), monkeys first learn to discriminate conditioned cues predicting an aversive puff of air (AX) from conditioned cues signaling safety (BX), and are then presented with the AB summation test. Similar to healthy rodents and humans, monkeys discriminate between A and B cues and show an attenuated startle response when presented with the AB compound (Fig. 2). After the summation tests, animals receive extinction trials in which A is presented without the aversive air puff. This powerful paradigm allows safety signal learning and fear extinction to be studied within the same subject.

A developmental lesion study was conducted to examine the long-term effects of early damage to the amygdala, hippocampus, or the orbital frontal cortex (Brodmann areas 11 and 13). Surprisingly, safety signal learning and fear extinction were intact despite early damage to these structures (Kazama et al., 2012). Presently, we speculate that compensatory mechanisms occurred across development, which have been reported in other tasks (Glavis-Bloom et al., 2008). Additionally, although area 13 of the orbital frontal cortex has been shown to share strong connectivity with the amygdala (for review, see Barbas, 2007; Price, 2007), human neuroimaging suggests that medial areas such as Brodmann area 14 may be more involved in flexibly modulating fear (Milad et al., 2009; Schuff et al., 2011). Additionally, neurophysiological data from nonhuman primate models suggest that dopaminergic striatal neurons may provide crucial prediction error feedback in the presence of safety signals (Matsumoto and Hikosaka, 2009). Thus, safety signal processing likely involves multiple brain areas that have only begun to be examined. Much work remains but the nonhuman primate AX+/BX- paradigm will likely permit tremendous advances as a translational tool.

**Safety signals as a reinforcer**

Safety signals provide relief from fearful states and may reinforce safety-seeking behavior. The relief experienced during the presentation of a safety signal may also motivate avoidance behavior symptomatic of many anxiety disorders. This possibility is relevant to anxiety disorders such as obsessive-compulsive disorder. For example, patients with obsessional checking rituals have reported a feeling of relief following the completion of checking behaviors (Roper et al., 1973). The reinforcing properties of a conditioned inhibitor of fear can be understood conceptually

when considering the proposal of Konorski (1967) of two opposing motivational systems that reciprocally inhibit one another, an aversive system and an appetitive system. He and others (Dickinson and Pearce, 1977; Dickinson and Dearing, 1979; Gray 1987) argued that inhibition of the aversive system by a safety signal should disinhibit the appetitive system and would therefore be functionally (or behaviorally) equivalent to a direct excitator of the appetitive system. A. B. P. Fernando, G. P. Urcelay, A. C. Mar, A. D. Dickinson, and T. W. Robbins (unpublished data) tested this hypothesis in rats by training a pavlovian conditioned inhibitor, an auditory stimulus presented in the absence of a mild footshock in one group, and a pavlovian conditioned appetitive excitator, an auditory stimulus paired with a sucrose pellet in another group. They hypothesized that if relief functions as reward, it should support the acquisition of a new behavioral response with subjects responding purely for the presentation of the inhibitor in the absence of any footshock (A.B.P. Fernando, personal communication). The appetitive stimulus preferentially supported a new instrumental response that was potentiated with D-amphetamine, whereas the safety signal did not support a new instrumental response and therefore did not demonstrate reinforcing properties.

Safety signals can also be provided contingent upon performance of an escape or avoidance response (Soltysik and Zielinski, 1962; Weisman and Litner, 1966, 1969a,b, 1971; Weisman et al., 1966; Dinsmoor, 2001). Such stimuli become associated with the relief from shock and should reinforce the avoidance response. Rats were trained on an instrumental lever press avoidance task to see whether an instrumentally trained safety signal would reinforce an already acquired avoidance response (A.B.P. Fernando G. P. Urcelay, A. C. Mar, A. D. Dickinson, and T. W. Robbins, personal communication). Akin to Rescorla (1969a), rats preferentially responded in a two-lever avoidance task on the lever that both prevented shock and produced the safety signal. Rats continued to demonstrate this preferential responding on the lever producing the safety signal in extinction. Both tests provide clear evidence that an instrumentally trained safety signal possesses reinforcing properties. Despite the demonstration of this study and others that safety signals possess reinforcing properties, there are clearly multiple lines of information processing associated with safety signals.

Safety signals alert the organism to when the environment is safe, thus promoting behaviors leading to natural rewards, such as feeding and mating, whereas danger signals inhibit these behaviors. One example that highlights how these antagonistic motivational systems can have bidirectional effects on behavior can be seen in rats that suppress lever pressing for food when a danger cue is presented but increase lever-pressing for food when a safety cue is presented (Walasek et al., 1995), demonstrating that safety signals can promote food-seeking behavior. Interestingly, even *Drosophila* display a similar antagonistic avoidance–approach behavior, showing conditioned avoidance of a danger odor and conditioned approach to the same odor cue if it signifies safety (Tanimoto et al., 2004).

Normally, obtaining natural rewards involves activating exploratory behavior, and several safety-conditioning protocols have demonstrated increases in exploratory behavior. Safety cues can increase exploration in an environment that is normally anxiogenic to mice, demonstrating that safety cues can take on anxiolytic properties and can even be used to condition a place preference (Rogan et al., 2005; Pollak et al., 2008). Distinguishing between the fear-inhibiting versus rewarding qualities of a safety signal is not simple, and the behavioral phenomena suggest over-

lap in the neural circuitry. Sangha et al. (S. Sangha, J. Z. Chadick, and P. H. Janak, unpublished data) developed a safety signal-training protocol that allows the parallel investigation of fear, safety, and reward learning in rats with simultaneous single unit recordings in the BA of freely behaving rats. As the rats flexibly switched among freezing, inhibition of freezing, and reward seeking, several populations of neurons emerged. As expected, many neurons changed (i.e., either increased or decreased) their firing preferentially to either the danger cue or the reward cue. Importantly, a population of neurons showed selective responses to the safety cue. Finally, there was another population of neurons in the BA that showed a similar change in firing rate in response to the safety cue and reward cue, implying that there is an overlap of safety and reward encoding in the BA. These data are the first to correlate amygdala single-unit activity with safety signals and behavior *in vivo*. In addition to Konorski's 1967 proposal of the safety signal disinhibiting the appetitive system through its inhibition of the aversive system, the observed overlap in neural encoding of safety and reward cues suggests that a safety signal may directly activate the appetitive system.

## Conclusion

Conditioned inhibition of fear is not a new concept to neuroscience; however, it is surprising that in 85 years since Pavlov's seminal work very little is known regarding the neural mechanisms underlying this phenomenon. This may come as a surprise to those familiar with fear conditioning and fear extinction for which tremendous advances, from the anatomical tracts to the molecular cascades, have been made. One reason for the disparity is that there is no standardized approach to train a safety signal, and the ones typically used are complex and do not lend themselves as easily to modern neuroscience approaches as pavlovian fear conditioning.

As we prepared the Mini-Symposium panel and this article, our goal was to bring safety learning to a broad audience and identify the large gaps in our current understanding. Despite the gaps, however, it is possible to suggest a framework for continued investigation of safety learning. We have addressed three components of safety signal processing: learning, fear inhibition in a summation test, and rewarding characteristics. First, safety-learning protocols require learning of a danger CS and then that a safety stimulus is associated with the nonoccurrence of danger. Such discrimination might involve generation of an expectation error signal and subsequent updating of predictive value of the safety signal. Numerous structures including dorsal striatum, PFC, periaqueductal gray, and amygdala have been implicated in expectation errors (Schultz and Dickinson, 2000; Belova et al., 2007; McNally et al., 2011); therefore, these structures are potential contributors to safety learning. Second, tests for safety learning involve flexible behavioral responses that switch as danger and safety cues are presented to the subject. This process depends on recall of the learned cues; the putative anatomical loci for storage of the safety signal are unknown. Behavioral flexibility depends upon contributions from vmPFC, orbital frontal cortex, and the striatum (Murray and Izquierdo, 2007; Wolfensteller and Ruge, 2012). Again, interactions of these structures with the amygdala would be required to switch behavior from avoidance to approach as danger and safety signals appear in the environment. Third, safety signals possess rewarding properties. Like the fear circuitry, reward circuitry is well understood involving the ventral tegmental area to the nucleus accumbens dopaminergic circuit (Koob, 1992), and interactions between these structures and the amygdala are important for goal-

directed behaviors (Schoenbaum et al., 1998, 2003). Thus, multiple circuits must interact with the amygdala to acquire, recall, and use safety signals.

Future studies of safety signal processing must account for the different roles of the amygdala nuclei and cell populations known to be involved in both danger and safety learning (Ehrlich et al., 2009; Ostroff et al., 2010; Christianson et al., 2011). Fortunately, technologies including *in vivo* electrophysiology, optogenetics (Tye et al., 2011), designer receptors exclusively activated by designer drugs (or DREADD) (Dong et al., 2010), and genetic tools allow for neural observations and manipulations that are both anatomically and temporally precise—a requirement for dissecting safety from danger. Furthermore, Pollak et al. (2008, 2010a,b) established a safety learning protocol that is translatable between mouse models and humans (Pollak et al., 2010a), providing the field with yet another powerful translational research tool (Pollak et al., 2010b). Equipped with a modern neuroscience toolkit and an important clinical correlate, we expect the next 85 years will welcome many exciting developments in our understanding of safety signals.

## References

- Amano T, Unal CT, Paré D (2010) Synaptic correlates of fear extinction in the amygdala. *Nat Neurosci* 13:489–494. [CrossRef Medline](#)
- Barbas H (2007) Flow of information for emotions through temporal and orbitofrontal pathways. *J Anat* 211:237–249. [CrossRef Medline](#)
- Belova MA, Paton JJ, Morrison SE, Salzman CD (2007) Expectation modulates neural responses to pleasant and aversive stimuli in primate amygdala. *Neuron* 55:970–984. [CrossRef Medline](#)
- Bouton ME (2004) Context and behavioral processes in extinction. *Learn Mem* 11:485–494. [CrossRef Medline](#)
- Christianson JP, Benison AM, Jennings J, Sandsmark EK, Amat J, Kaufman RD, Baratta MV, Paul ED, Campeau S, Watkins LR, Barth DS, Maier SF (2008) The sensory insular cortex mediates the stress-buffering effects of safety signals but not behavioral control. *J Neurosci* 28:13703–13711. [CrossRef Medline](#)
- Christianson JP, Jennings JH, Ragole T, Flyer JG, Benison AM, Barth DS, Watkins LR, Maier SF (2011) Safety signals mitigate the consequences of uncontrollable stress via a circuit involving the sensory insular cortex and bed nucleus of the stria terminalis. *Biol Psychiatry* 70:458–464. [CrossRef Medline](#)
- Corbo V, Clément MH, Armony JL, Pruessner JC, Brunet A (2005) Size versus shape differences: contrasting voxel-based and volumetric analyses of the anterior cingulate cortex in individuals with acute posttraumatic stress disorder. *Biol Psychiatry* 58:119–124. [CrossRef Medline](#)
- Dickinson A, Dearing MF (1979) Appetitive-aversive interactions and inhibitory processes. In: *Mechanism of learning and motivation* (Dickinson A, Boakes RA, eds.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Dickinson A, Pearce JM (1977) Inhibitory interactions between appetitive and aversive stimuli. *Psychol Bull* 84:690–711. [CrossRef](#)
- Dinsmoor JA (2001) Stimuli inevitably generated by behavior that avoids electric shock are inherently reinforcing. *J Exp Anal Behav* 75:311–333. [CrossRef Medline](#)
- Dong S, Allen JA, Farrell M, Roth BL (2010) A chemical-genetic approach for precise spatio-temporal control of cellular signaling. *Mol Biosyst* 6:1376–1380. [CrossRef Medline](#)
- Ehrlich I, Humeau Y, Grenier F, Ciochi S, Herry C, Lüthi A (2009) Amygdala inhibitory circuits and the control of fear memory. *Neuron* 62:757–771. [CrossRef Medline](#)
- Etkin A, Wager TD (2007) Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *Am J Psychiatry* 164:1476–1488. [CrossRef Medline](#)
- Falls WA, Davis M (1995) Lesions of the central nucleus of the amygdala block conditioned excitation, but not conditioned inhibition of fear as measured with the fear-potentiated startle effect. *Behav Neurosci* 109:379–387. [CrossRef Medline](#)
- Falls WA, Bakken KT, Heldt SA (1997) Lesions of the perirhinal cortex interfere with conditioned excitation but not with conditioned inhibition of fear. *Behav Neurosci* 111:476–486. [CrossRef Medline](#)
- Gewirtz JC, Falls WA, Davis M (1997) Normal conditioned inhibition and extinction of freezing and fear-potentiated startle following electrolytic lesions of medial prefrontal cortex in rats. *Behav Neurosci* 111:712–726. [CrossRef Medline](#)
- Glavis-Bloom CK, Kazama AM, Bachevalier J (2008) Paradoxical functional facilitation in rhesus macaques with neonatal hippocampal lesions on two stimulus-reward association tasks. *Soc Neurosci Abstr* 34:791.1.
- Gray JA (1987) *The psychology of fear and stress*, Ed 2. Cambridge, UK: Cambridge UP.
- Hammond LJ (1967) A traditional demonstration of the properties of Pavlovian inhibition using differential CER. *Psychon Sci* 9:65–66.
- Heldt SA, Falls WA (2006) Posttraining lesions of the auditory thalamus, but not cortex, disrupt the inhibition of fear conditioned to an auditory stimulus. *Eur J Neurosci* 23:765–779. [CrossRef Medline](#)
- Heldt SA, Coover GD, Falls WA (2002) Posttraining but not pretraining lesions of the hippocampus interfere with feature-negative discrimination of fear-potentiated startle. *Hippocampus* 12:774–786. [CrossRef Medline](#)
- Hughes KC, Shin LM (2011) Functional neuroimaging studies of post-traumatic stress disorder. *Expert Rev Neurother* 11:275–285. [CrossRef Medline](#)
- Josselyn SA, Falls WA, Gewirtz JC, Pistell P, Davis M (2005) The nucleus accumbens is not critically involved in mediating the effects of a safety signal on behavior. *Neuropsychopharmacology* 30:17–26. [CrossRef Medline](#)
- Jovanovic T, Keyes M, Fiallos A, Myers KM, Davis M, Duncan EJ (2005) Fear potentiation and fear inhibition in a human fear-potentiated startle paradigm. *Biol Psychiatry* 57:1559–1564. [CrossRef Medline](#)
- Jovanovic T, Norrholm SD, Fennell JE, Keyes M, Fiallos AM, Myers KM, Davis M, Duncan EJ (2009) Posttraumatic stress disorder may be associated with impaired fear inhibition: relation to symptom severity. *Psychiatry Res* 167:151–160. [CrossRef Medline](#)
- Jovanovic T, Norrholm SD, Blanding NQ, Davis M, Duncan E, Bradley B, Ressler KJ (2010) Impaired fear inhibition is a biomarker of PTSD but not depression. *Depress Anxiety* 27:244–251. [CrossRef Medline](#)
- Jovanovic T, Kazama A, Bachevalier J, Davis M (2012a) Impaired safety signal learning may be a biomarker of PTSD. *Neuropharmacology* 62:695–704. [CrossRef Medline](#)
- Jovanovic T, Sakoman AJ, Kozarić-Kovačić D, Meštrović AH, Duncan EJ, Davis M, Norrholm SD (2012b) Acute stress disorder versus chronic posttraumatic stress disorder: inhibition of fear as a function of time since trauma. *Depress Anxiety*. Advance online publication. doi: 10.1002/da.21991.
- Kazama AM, Heuer E, Davis M, Bachevalier J (2012) Effects of neonatal amygdala lesions on fear learning, conditioned inhibition, and extinction in adult macaques. *Behav Neurosci* 126:392–403. [CrossRef Medline](#)
- Kim JJ, Jung MW (2006) Neural circuits and mechanisms involved in Pavlovian fear conditioning: a critical review. *Neurosci Biobehav Rev* 30:188–202. [CrossRef Medline](#)
- Konorski J (1948) *Conditioned reflexes and neuron organization*. Cambridge, UK: Cambridge UP.
- Konorski J (1967) *Integrative activity of the brain*. Chicago: University of Chicago.
- Koob GF (1992) Drugs of abuse: anatomy, pharmacology and function of reward pathways. *Trends Pharmacol Sci* 13:177–184. [CrossRef Medline](#)
- Maier SF, Grahn RE, Kalman BA, Sutton LC, Wiertelak EP, Watkins LR (1993) The role of the amygdala and dorsal raphe nucleus in mediating the behavioral consequences of inescapable shock. *Behav Neurosci* 107:377–388. [CrossRef Medline](#)
- Matsumoto M, Hikosaka O (2009) Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* 459:837–841. [CrossRef Medline](#)
- McDonald AJ, Shammah-Lagnado SJ, Shi C, Davis M (1999) Cortical afferents to the extended amygdala. *Ann N Y Acad Sci* 877:309–338. [CrossRef Medline](#)
- McNally GP, Johansen JP, Blair HT (2011) Placing prediction into the fear circuit. *Trends Neurosci* 34:283–292. [CrossRef Medline](#)
- Milad MR, Quirk GJ (2002) Neurons in medial prefrontal cortex signal memory for fear extinction. *Nature* 420:70–74. [CrossRef Medline](#)
- Milad MR, Pitman RK, Ellis CB, Gold AL, Shin LM, Lasko NB, Zeidan MA, Handwerker K, Orr SP, Rauch SL (2009) Neurobiological basis of failure

- to recall extinction memory in posttraumatic stress disorder. *Biol Psychiatry* 66:1075–1082. [CrossRef Medline](#)
- Miller RR, Spear NS (Eds.) (1985) Information processing in animals: conditioned inhibition. Hillsdale, NJ: Erlbaum.
- Mineka S, Cook M, Miller S (1984) Fear conditioned with escapable and inescapable shock: effects of a feedback stimulus. *J Exp Psychol Anim Behav Process* 10:307–323. [CrossRef](#)
- Murray EA, Izquierdo A (2007) Orbitofrontal cortex and amygdala contributions to affect and action in primates. *Ann N Y Acad Sci* 1121:273–296. [CrossRef Medline](#)
- Myers KM, Davis M (2004) AX+, BX- discrimination learning in the fear-potentiated startle paradigm: possible relevance to inhibitory fear learning in extinction. *Learn Mem* 11:464–475. [CrossRef Medline](#)
- Nieuwenhuis R (2012) The insular cortex: a review. *Prog Brain Res* 195:123–163. [CrossRef Medline](#)
- Ongür D, Price JL (2000) The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb Cortex* 10:206–219. [CrossRef Medline](#)
- Ostroff LE, Cain CK, Bedont J, Monfils MH, Ledoux JE (2010) Fear and safety learning differentially affect synapse size and dendritic translation in the lateral amygdala. *Proc Natl Acad Sci U S A* 107:9418–9423. [CrossRef Medline](#)
- Ostroff LE, Cain CK, Jindal N, Dar N, Ledoux JE (2012) Stability of presynaptic vesicle pools and changes in synapse morphology in the amygdala following fear learning in adult rats. *J Comp Neurol* 520:295–314. [CrossRef Medline](#)
- Paulus MP, Stein MB (2006) An insular view of anxiety. *Biol Psychiatry* 60:383–387. [CrossRef Medline](#)
- Pavlov IP (1927) Conditioned reflexes. London: Oxford UP.
- Pollak DD, Monje FJ, Zuckerman L, Denny CA, Drew MR, Kandel ER (2008) An animal model of a behavioral intervention for depression. *Neuron* 60:149–161. [CrossRef Medline](#)
- Pollak DD, Monje FJ, Lubec G (2010a) The learned safety paradigm as a mouse model for neuropsychiatric research. *Nat Protoc* 5:954–962. [CrossRef Medline](#)
- Pollak DD, Rogan MT, Egner T, Perez DL, Yanagihara TK, Hirsch J (2010b) A translational bridge between mouse and human models of learned safety. *Ann Med* 42:115–122. [CrossRef Medline](#)
- Price JL (2007) Definition of the orbital cortex in relation to specific connections with limbic and visceral structures and other cortical regions. *Ann N Y Acad Sci* 1121:54–71. [CrossRef Medline](#)
- Rauch SL, Shin LM, Phelps EA (2006) Neurocircuitry models of posttraumatic stress disorder and extinction: human neuroimaging research—past, present, and future. *Biol Psychiatry* 60:376–382. [CrossRef Medline](#)
- Rescorla RA (1969a) Establishment of a positive reinforcer through contrast with shock. *J Comp Physiol Psychol* 67:260–263. [CrossRef Medline](#)
- Rescorla RA (1969b) Pavlovian conditioned inhibition. *Psychol Bull* 72:77–94. [CrossRef](#)
- Rodgers KM, Benison AM, Klein A, Barth DS (2008) Auditory, somatosensory, and multisensory insular cortex in the rat. *Cereb Cortex* 18:2941–2951. [CrossRef Medline](#)
- Rogan MT, Stäubli UV, LeDoux JE (1997) Fear conditioning induces associative long-term potentiation in the amygdala. *Nature* 390:604–607. [CrossRef Medline](#)
- Rogan MT, Leon KS, Perez DL, Kandel ER (2005) Distinct neural signatures for safety and danger in the amygdala and striatum of the mouse. *Neuron* 46:309–320. [CrossRef Medline](#)
- Roper G, Rachman S, Hodgson R (1973) An experiment on obsessional checking. *Behav Res Ther* 11:271–277. [CrossRef Medline](#)
- Sah P, Faber ES, Lopez De Armentia M, Power J (2003) The amygdaloid complex: anatomy and physiology. *Physiol Rev* 83:803–834. [CrossRef Medline](#)
- Schoenbaum G, Chiba AA, Gallagher M (1998) Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nat Neurosci* 1:155–159. [CrossRef Medline](#)
- Schoenbaum G, Setlow B, Saddoris MP, Gallagher M (2003) Encoding predicted outcome and acquired value in orbitofrontal cortex during cue sampling depends upon input from basolateral amygdala. *Neuron* 39:855–867. [CrossRef Medline](#)
- Schuff N, Zhang Y, Zhan W, Lenoci M, Ching C, Boreta L, Mueller SG, Wang Z, Marmar CR, Weiner MW, Neylan TC (2011) Patterns of altered cortical perfusion and diminished subcortical integrity in posttraumatic stress disorder: an MRI study. *Neuroimage* 54:S62–S68. [CrossRef Medline](#)
- Schultz W, Dickinson A (2000) Neuronal coding of prediction errors. *Annu Rev Neurosci* 23:473–500. [CrossRef Medline](#)
- Soltysik S, Zielinski K (1962) Conditioned inhibition of the avoidance reflex. *Acta Biol Exp (Warsz)* 22:157–167.
- Tanimoto H, Heisenberg M, Gerber B (2004) Event timing turns punishment to reward. *Nature* 430:983. [CrossRef Medline](#)
- Tronson NC, Corcoran KA, Jovasevic V, Radulovic J (2012) Fear conditioning and extinction: emotional states encoded by distinct signaling pathways. *Trends Neurosci* 35:145–155. [CrossRef Medline](#)
- Tye KM, Prakash R, Kim SY, Fenno LE, Grosenick L, Zarabi H, Thompson KR, Gradinaru V, Ramakrishnan C, Deisseroth K (2011) Amygdala circuitry mediating reversible and bidirectional control of anxiety. *Nature* 471:358–362. [CrossRef Medline](#)
- Waddell J, Heldt S, Falls WA (2003) Posttraining lesion of the superior colliculus interferes with feature-negative discrimination of fear-potentiated startle. *Behav Brain Res* 142:115–124. [CrossRef Medline](#)
- Walasek G, Wesierska M, Zielinski K (1995) Conditioning of fear and conditioning of safety in rats. *Acta Neurobiol Exp* 55:121–132.
- Watkins LR, Wiertelak EP, McGorry M, Martinez J, Schwartz B, Sisk D, Maier SF (1998) Neurocircuitry of conditioned inhibition of analgesia: effects of amygdala, dorsal raphe, ventral medullary, and spinal cord lesions on antianalgesia in the rat. *Behav Neurosci* 112:360–378. [CrossRef Medline](#)
- Weisman RG, Litner JS (1969a) The course of Pavlovian excitation and inhibition of fear in rats. *J Comp Physiol Psychol* 69:667–672. [CrossRef Medline](#)
- Weisman RG, Litner JS (1969b) Positive conditioned reinforcement of Sidman avoidance behavior in rats. *J Comp Physiol Psychol* 68:597–603. [CrossRef](#)
- Weisman RG, Litner JS (1971) Role of the intertrial interval in Pavlovian differential conditioning of fear in rats. *J Comp Physiol Psychol* 74:211–218. [CrossRef Medline](#)
- Weisman RG, Denny MR, Platt SA, Zerbolio DJ Jr (1966) Facilitation of extinction by a stimulus associated with long nonshock confinement periods. *J Comp Physiol Psychol* 62:26–30. [CrossRef Medline](#)
- Weiss JM (1971) Effects of coping behavior with and without a feedback signal on stress pathology in rats. *J Comp Physiol Psychol* 77:22–30. [CrossRef Medline](#)
- Wiertelak EP, Maier SF, Watkins LR (1992) Cholecystokinin antianalgesia: safety cues abolish morphine analgesia. *Science* 256:830–833. [CrossRef Medline](#)
- Winslow JT, Noble PL, Davis M (2008) AX+/BX- discrimination learning in the fear-potentiated startle paradigm in monkeys. *Learn Mem* 15:63–66. [CrossRef Medline](#)
- Wolfensteller U, Ruge H (2012) Frontostriatal mechanisms in instruction-based learning as a hallmark of flexible goal-directed behavior. *Front Psychol* 3:192. [Medline](#)