A model of amygdala-hippocampal-prefrontal interaction in fear conditioning and extinction in animals

Ahmed A. Moustafa[1,2,3], **Mark W. Gilbertson**[4,5], **Scott P. Orr**[5,6], **Mohammad M. Herzallah**[2,7], **Richard. J. Servatius**[3,8], and **Catherine E. Myers**[3,8,9]

[1]School of Social Sciences and Psychology & Marcs Institute for Brain and Behaviour, University of Western Sydney, Sydney, NSW, Australia.

[2]Center for Molecular and Behavioral Neuroscience, Rutgers University, Newark, NJ.

[3]Department of Veterans Affairs, New Jersey Health Care System, East Orange, NJ.

[4]Manchester VA Medical Center, Manchester, NH.

[5]Harvard Medical School, Boston, MA

[6]Massachusetts General Hospital, Boston, MA.

[7]Cognitive Neuroscience Lab, Al Quds University-Abu Dis, Palestinian Territories.

[8]Stress & Motivated Behavior Institute, New Jersey Medical School, University of Medicine and Dentistry of New Jersey, Newark, NJ.

[9]Department of Psychology, Rutgers University-Newark, Newark, NJ.

Abstract

Empirical research has shown that the amygdala, hippocampus, and ventromedial prefrontal cortex (vmPFC) are involved in fear conditioning. However, the functional contribution of each brain area and the nature of their interactions are not clearly understood. Here, we extend existing neural network models of the functional roles of the hippocampus in classical conditioning to include interactions with the amygdala and prefrontal cortex. We apply the model to fear conditioning, in which animals learn physiological (e.g. heart rate) and behavioral (e.g. freezing) responses to stimuli that have been paired with a highly aversive event (e.g. electrical shock). The key feature of our model is that learning of these conditioned responses in the central nucleus of the amygdala is modulated by two separate processes, one from basolateral amygdala and signaling a positive prediction error, and one from the vmPFC, via the intercalated cells of the amygdala, and signaling a negative prediction error. In addition, we propose that hippocampal input to both vmPFC and basolateral amygdala is essential for contextual modulation of fear acquisition and extinction. The model is sufficient to account for a body of data from various animal fear conditioning paradigms, including acquisition, extinction, reacquisition, and context specificity effects. Consistent with studies on lesioned animals, our model shows that damage to the vmPFC impairs extinction, while damage to the hippocampus impairs extinction in a different context (e.g., a different conditioning chamber from that used in initial training in animal experiments).

To whom correspondence should be addressed, **Ahmed A. Moustafa**, School of Psychology, University of Western Sydney, Sydney, NSW, Australia. a.moustafa@uws.edu.au.

We also discuss model limitations and predictions, including the effects of number of training trials on fear conditioning.

Keywords

fear conditioning; computational model; hippocampus; amygdala; ventromedial prefrontal cortex; extinction

1-Introduction

Several brain structures – including the amygdala, hippocampus, and ventromedial prefrontal cortex (vmPFC) – are involved in fear conditioning and extinction, yet the functional contribution of each brain area and the nature of their interactions are not clearly understood. Here, we propose a neural network model that addresses the function of these brain areas in fear conditioning and extinction.

In classical conditioning, a previously-neutral stimulus (the conditioned stimulus or CS) is repeatedly paired with an unconditioned stimulus (US) that evokes a reflexive response. Over time, the CS itself can come to evoke an anticipatory response, the conditioned response or CR. In somatomotor conditioning, the US may be an airpuff or mild periorbital shock that evokes a protective eye closure; the CR is then an anticipatory eyeblink, so that the eye is partially closed at the time of expected US arrival. Prior work has shown that the cerebellum is the necessary and sufficient substrate for eyeblink conditioning (Christian & Thompson, 2003), but that other areas including the hippocampus may also participate and may even be critical, depending on the nature and timing of the various stimuli.

Other kinds of classical conditioning depend on other brain substrates. For example, fear conditioning in animals refers to a broad class of paradigms in which the CS is paired with an aversive stimulus, such as electric shock; with repeated pairings, the CS can come to evoke a range of "fear responses," including physiological responses (changes in heart rate, blood pressure, etc.), and behavioral responses (freezing, startle, etc.). The choice of CR to be measured in a particular experiment depends on many factors including the species under study. For example, heart rate conditioning has been used in rabbits (Gibbs, Prescott, & Powell, 1992; Kapp, Frysinger, Gallagher, & Haselton, 1979), which typically show a decrease in heart rate (e.g., bradycardia) in response to a CS paired with shock. In contrast, a large body of research on rats has considered the freezing CR, a brief cessation of ongoing behavior in response to a CS paired with shock (Duvarci, Popa, & Pare, 2011; LeDoux, 1992). Fear conditioning shares many features with somatomotor conditioning, including a negatively accelerated learning curve, extinction of learned responses when the CS no longer predicts the US, and sensitivity to context. However, there are also differences: for example, whereas eyeblink CRs are typically acquired following hundreds of CS-US pairings in the rabbit, fear responses such as heart-rate changes and freezing can be acquired within a few (or even a single) CS-US pairings.

Below, we first review literature on behavioral results obtained from fear conditioning paradigms in animals, then review the known neural substrates of this learning, and finally discuss how the model proposes these brain substrates interact during fear conditioning. We then describe the model simulations, and present simulation data showing that this model of amygdala-hippocampal-prefrontal interaction provides a qualitative fit to a range of data obtained from animal fear conditioning studies.

1-1 Relevant empirical background: behavioral paradigms and neural substrates

Below, we discuss behavioral paradigms of fear conditioning, and then discuss neural substrates of fear conditioning.

1-1-1 Behavioral paradigms of fear conditioning—Whereas acquisition involves increased expression of the CR as a result of CS-US pairing, extinction refers to the reduction of CR expression when the CS is no longer paired with the US. As with other kinds of learning, extinction of fear CRs could be a consequence of either erasing previously acquired fear memories or forming inhibitory fear responses that overcome or compete with previously acquired fear responses. Most studies have shown that fear extinction involves forming new extinction memories that overrule the previously-acquired fear response (Bouton, Westbrook, Corcoran, & Maren, 2006; Milad & Quirk, 2002; Myers & Davis, 2007).

In addition to the acquisition and extinction of fear, other conditioning procedures have also examined renewal and reacquisition (Bouton & King, 1983; Herry et al., 2008; Hobin, Ji, & Maren, 2006; Ji & Maren, 2007; Milad, Orr, Pitman, & Rauch, 2005; Zelikowsky, Pham, & Fanselow, 2011). The reacquisition paradigm involves three phases. In the first phase, all subjects are trained to acquire a fear response in one context (e.g., a training chamber with certain features and odors). In the second phase, all subjects are given extinction trials in a different context (e.g., a different conditioning chamber with different features and odors). In the last phase, half of the subjects group are trained on fear acquisition using the acquisition context (i.e. as in the first phase), and the other half of subjects are trained on fear acquisition using the extinction context (i.e., as in the second phase) (Bouton & King, 1983). Interestingly, studies have found that reacquiring fear responses is faster than in the initial fear acquisition phase (Leung, Bailey, Laurent, & Westbrook, 2007). Other fear conditioning procedures have examined the effect of changing the background context during the extinction or reacquisition phase (Bouton, 1984; Bouton & King, 1983; Bouton & Peck, 1989; Corcoran & Maren, 2001). In experimental settings, context usually refers to spatial and olfactory features of the testing box and/or other external cues that might have been used by the subjects during learning. Studies have shown that extinction learning is quicker when it occurs in a different context than that used in the fear acquisition phase (Corcoran & Maren, 2001; for discussion see Delamater, 2004). Similarly, reacquisition in the third phase is quicker when using the acquisition rather than extinction context (Bouton & King, 1983; Milad, et al., 2005). In sum, the goal of this project is to link all these behavioral paradigms in one coherent framework. Table 1 describes these fear conditioning tasks.

1-1-2 Neural substrates of fear conditioning—Various lesion, physiological, and imaging studies have investigated the neural basis of fear conditioning in rats and rabbits, as well as other animals. Most of these studies have found that three different brain areas have been implicated in fear conditioning: amygdala, hippocampus, and vmPFC. Figure 1 shows a simplified anatomical map of their connectivity. Research on fear conditioning has been attempting to elucidate the specific contribution of each brain area to fear conditioning. Below, we review empirical studies on the role of amygdala, hippocampus, and vmPFC in fear conditioning and extinction (for a review, see Jovanovic & Ressler; Maren & Quirk, 2004).

<u>Amygdala:</u> The amygdala is a collection of highly differentiated nuclei that belong to different functional systems (Swanson & Petrovich, 1998). One subregion of the amygdala is the central nucleus (CeA), a part of the motor striatum with mainly GABAergic projections to autonomic systems. Another major subregion of the amygdala is the lateral

and anterior basolateral nuclei (BLA), which may be considered ventromedial extensions of the temporal and frontal lobes with mainly glutamatergic projections (Swanson & Petrovich, 1998).

The central nucleus of the amygdala (CeA) is involved in the initiation of various fear responses, including freezing and heart rate changes (Duvarci, et al., 2011). Conditioned freezing and heart rate changes appear to be driven by outputs from CeA through the lateral hypothalamus to the cardioinhibitory neurons in the dorsal vagal nucleus and/or the nucleus of the solitary tract (McCabe, Gentile, Markgraf, Teich, & Schneiderman, 1992; Wiersma, Bohus, & Koolhaas, 1993). Specifically, the CeA projects to the parasympathetic nervous system (driving freezing, heart rate changes, respiratory changes), the hypothalamus (driving the release of stress hormones), and brainstem motor areas (driving motor responses such as freezing). Stimulation of CeA can produce altered heart rate (Cox et al., 1986), and neurons in the CeA show CS-evoked learning-related responses during both heart rate and eyeblink conditioning (e.g. Rorick-Kehn & Steinmetz, 2005); these responses are correlated with the magnitude of the CR in heart rate conditioning (Pascoe & Kapp, 1985). In freezing experiments, lesioning CeA abolishes fear responses (Davis, 1992; LeDoux, 1992; McCabe, et al., 1992). Amygdala lesion (specifically CeA lesion) also abolishes heart rate conditioning (Kapp, et al., 1979; McCabe, et al., 1992; McCabe, McEchron, Green, & Schneiderman, 1993; Sananes & Campbell, 1989; Weisz, Harden, & Xiang, 1992).

Anatomical studies have shown that the CeA receives projections from the BLA and intercalated cells (ITC), which are responsible for initiating and inhibiting fear responses (Amano, Unal, & Pare, 2010). The BLA is a key site of sensory integration in the amygdala (Aggleton, 2000; LeDoux, 1993). Neurons in the BLA change sensory-evoked responses to stimuli that have been paired with reinforcement (see e.g. Corbit & Balleine, 2005). *In vivo* studies have documented learning-related changes, in the form of long-term plasticity (Rosenkranz & Grace, 2002) in the BLA during classical conditioning of an odor CS paired with a footshock US.

The nuclei of the BLA project, directly or indirectly, to CeA (Swanson & Petrovich, 1998). Somatosensory shock information may be conveyed to BLA from thalamic nuclei (Sah, Faber, Lopez De Armentia, & Power, 2003) and possibly also via routes not synapsing in thalamus (Lanuza, Nader, & Ledoux, 2004). Herry et al. (2008) have recorded from basal amygdala neurons while mice are trained on a fear conditioning and renewal paradigm. Interestingly, they have found that some lateral amygdala neurons increase their activation during fear acquisition, but not extinction. These neurons also increased their activation levels during the renewal phase; this supports the view that BLA is essential for forming fear acquisition memories.

Research has shown that the amygdala is involved in both fear acquisition and extinction (Indovina, Robbins, Nunez-Elizalde, Dunn, & Bishop, 2011; for review see Kim & Jung, 2006; LaBar, Gatenby, Gore, LeDoux, & Phelps, 1998; LeDoux, 1993; Parkes & Westbrook, 2010; Phelps, Delgado, Nearing, & LeDoux, 2004). Physiological studies have shown that different classes of neurons are involved in both fear acquisition and extinction (Amano, et al., 2010; Herry, et al., 2008): while basolateral neurons participate in fear acquisition, intercalated neurons play an important role in fear extinction.

Early studies have shown that lesioning BLA interferes with acquisition of fear responses (LeDoux, 1993; Maren, Aharonov, & Fanselow, 1996), while lesioning the CeA interferes with initiation of fear responses (for discussion see Bellgowan & Helmstetter, 1996; Kapp, et al., 1979; LeDoux, Iwata, Cicchetti, & Reis, 1988). Lesioning the BLA does not interfere with motor activity or reactions to shocks (Maren, 1998). Interestingly, Anglada-Figueroa

and Quirk (2005) have shown that basal amygdala lesion interferes with fear acquisition but not extinction. Similarly, Tronson, Wiseman, Olausson, and Taylor (2006) found that interfering with BLA function (by inhibiting amygdalar protein kinase A) does not affect fear extinction.

Besides the BLA, the CeA also receives projections from the ITC cells (Pare & Smith, 1993). The ITC cells are islands of GABAergic neurons that lie between the BLA and CeA, which are excited by vmPFC input, and which inhibit CeA neurons (Berretta, Pantazopoulos, Caldera, Pantazopoulos, & Pare, 2005; Likhtik, Pelletier, Paz, & Pare, 2005; Quirk & Gehlert; Quirk, Likhtik, Pelletier, & Pare, 2003). Also, ITC receives excitatory projections from the infralimbic cortex (Amir, Amano, & Pare, 2011), a pathway important for inhibiting fear responses (Amano, et al.). Physiological studies suggest that ITC neurons play an essential role in fear extinction (Manko, Geracitano, & Capogna, 2011; Pare & Smith, 1993). Lesion of ITC cells following extinction causes deficits in extinction expression, and the degree of CR expression is proportional to the number of surviving ITC cells (Likhtik, Popa, Apergis-Schoute, Fidacaro, & Pare, 2008).

Hippocampus: While the amygdala plays a key role in cue conditioning, the hippocampus is essential for learning to represent contextual information in fear conditioning in rats. Representing contextual information occurs when the subject learns not only about the explicit conditioned stimulus that warns of upcoming aversive events such as shock, but also about the context in which those aversive events occur. Typically, after exposure to a pairing of tone and shock in a particular experimental chamber, a rat will show a conditioned fear response (such as freezing) when the tone is presented later – but the rat will also show a fear response when returned to the chamber later (Anagnostaras, Maren, & Fanselow, 1999; Rudy, Huff, & Matus-Amat, 2004). Whereas conditioning to the cue depends on the amygdala, hippocampal lesion spares cue conditioning but interferes with conditioning to the context (Anagnostaras, et al., 1999; Kim & Fanselow, 1992; Maren & Fanselow, 1997; Phillips & LeDoux, 1992). It has been known since the 1960s that lesioning the hippocampus does not interfere with acquisition of fear CRs (Goldstein, 1965). Supporting a role for the hippocampal region in contextual encoding, animal studies have shown that different hippocampal cells respond to different contexts (Kentros, Agnihotri, Streater, Hawkins, & Kandel, 2004; Knierim, 2002; Leutgeb, Leutgeb, Moser, & Moser, 2007). The hippocampus is also important for contextual encoding of extinction memories (Ji & Maren, 2007).

Lesioning the hippocampus in rats abolishes context dependency/specificity in many paradigms, including context shift, extinction, reacquisition, renewal, and latent inhibition. For example, Ji and Maren (2005) found that hippocampal lesion abolishes context specificity in the renewal paradigm. Also, Fanselow and colleagues (Anagnostaras, et al., 1999) have found that hippocampal lesions impair contextual fear conditioning (where rats are shocked in an experimental chamber without the use of CSs) only if lesions are made one day but not more than 50 days after acquisition). Studies have also found that hippocampal lesion does not impair fear acquisition (Phillips & LeDoux, 1992), but impair extinction either in the same or new context (Corcoran & Maren, 2001; Ji & Maren, 2007). Similarly, hippocampal lesion also abolishes contextual effects in heart rate conditioning (Resstel, Joca, Correa, & Guimaraes, 2008).

Ventromedial prefrontal cortex (vmPFC): Anatomically, the vmPFC projects to the BLA and CeA in rabbits (Buchanan, Thompson, Maxwell, & Powell, 1994) and monkeys (Stefanacci & Amaral, 2002); infralimbic cortex in rats appears homologous to vmPFC (see Quirk & Mueller, 2008). Robust excitatory projections exist between hippocampus and vmPFC (Cavada, Company, Tejedor, Cruz-Rizzolo, & Reinoso-Suarez, 2000; Jay & Witter,

1991). Extensive studies have shown that the vmPFC is important for fear extinction in animals (Likhtik, et al., 2005; Milad & Quirk, 2002; Morgan, Romanski, & LeDoux; Quirk & Mueller, 2008). This line of research has shown that extinction involves the formation of new memories rather than *erasure* of older fear memories (for review see Myers & Davis, 2007). This has led researchers to introduce the term extinction memory, as separate from fear memory.

Infralimbic stimulation reduces the response rate of CeA neurons to input from lateral amygdala, suggesting that infralimbic cortex modulates the CeA-mediated expression of the CR (Pare, Quirk, & Ledoux, 2004; Quirk, et al., 2003). Infralimbic simulation facilitates extinction (Milad, Vidal-Gonzalez, & Quirk, 2004), suggesting it plays a role in inhibiting the expression of fear CRs (Milad & Quirk, 2002). Infralimbic neurons are particularly active on the day following extinction, and individual animals' retention of extinction correlates with infralimbic activity (Burgos-Robles, Vidal-Gonzalez, Santini, & Quirk, 2007; Milad & Quirk, 2002; Santini, Quirk, & Porter, 2008). Rats with infralimbic lesions condition and extinguish normally, but show increased spontaneous recovery when tested a day later (Quirk, Russo, Barron, & Lebron, 2000). Studies have also shown impaired fear extinction following infralimbic lesion (Lebron, Milad, & Quirk, 2004). Interestingly, empirical studies have found that applying microstimulation to cortical areas neighboring the infralimbic cortex, including prelimbic cortex and anterior cingulate, does not affect fear extinction (Sierra-Mercado, Corcoran, Lebron-Milad, & Quirk, 2006; Sierra-Mercado, Padilla-Coreano, & Quirk, 2010; Vidal-Gonzalez, Vidal-Gonzalez, Rauch, & Quirk, 2006).

1-2 Computational model: linking all together

How does the brain know when to encode either fear acquisition or extinction memory? This is important so that the BLA is recruited during fear acquisition, but the vmPFC is activated during fear extinction only. Research has shown that prediction of the US controls which brain areas are involved. For example, in humans given a classical conditioning task using visual CSs and mild electric shocks as the US, negative prediction errors produced by omission of an expected US were correlated with vmPFC activity (Spoormaker, Andrade, et al., 2011). In mice, negative prediction error during fear conditioning recruits different hippocampal cells from those involved in fear memory (Huh et al., 2009). We, here, argue that a positive prediction error enhances fear learning in the BLA, while a negative prediction error enhances extinction learning in the vmPFC (for empirical support see, Cole & McNally, 2009; McNally, Johansen, & Blair, 2011).

In our model, we simulate fear conditioning and extinction using the temporal difference (TD) algorithm. Positive TD error enhances learning in the BLA, while negative TD errors recruits and enhance learning in vmPFC. TD learning is an algorithm that learns CS-US associations (Sutton & Barto, 1998). The TD algorithm was first used by Witten (Witten, 1977) and Barto, Sutton, and Anderson (1983), but its roots can arguably be traced back to the animal learning theory field beginning with Thorndike (1911) and the theory of optimal control in engineering beginning with dynamic programming (Bellman, 1957). The TD model has been used to simulate rewarding (Montague, Dayan, & Sejnowski, 1996; Moustafa & Gluck, 2011) and aversive learning (Moutoussis, Bentall, Williams, & Dayan, 2008).

In sum, the main theory of our model is separating the effects of positive and negative prediction error, and treating them as separate processes inputting to BLA and vmPFC (see Discussion below for comparisons to other temporal difference models and models of extinction). In addition, in our model, we propose that the hippocampus provides a contextual signal that is projected to both the BLA and vmPFC, and is important for linking contextual with cue information during conditioning.

2- Methods

Below, we describe the details of our simulations of the hippocampal region, amygdala, and vmPFC. Our intent is not to simulate all of the anatomical and physiological details of these brain areas, but rather to build an initial framework in which to study systems-level interactions among these brain regions in fear conditioning. Model architecture is depicted in Figure 2. We train the model for 10 trials in each phase of the experiments described in Table 1. Results presented below are averaged across 50 runs of the model in each simulation study.

The input pattern consists of a 32-bit stimulus. This can specify the values of up to one CS, two contextual cues, and one US, that are manipulated depending on the simulation study (see Table 1). Each stimulus is represented using an 8-bit vector. As in our earlier models (A. A. Moustafa, Myers, & Gluck, 2009), in the input layer of this model, the representation of different stimuli or different contexts are orthogonal (i.e., different stimuli are encoded by different nodes).

All weights (i.e., connections among nodes in the network, which roughly correspond to biological synapses) in the model are initialized to random values in the range [0, 0.6]. All weights in the network (including simulated amygdala, hippocampus, and vmPFC) are perturbed using Gaussian noise (Moustafa & Maida, 2007). Let u_{ji} be the perturbed weight connecting unit i to unit j. Perturbed weight values are computed as follows:

$$u_{ji}(t) = w_{ji}(t) + \delta_{ji}(t),$$

where $\delta_{ji}(t)$ be the Gaussian noise associated with the weight w_{ji} (with zero mean and standard deviation of 0.025, also see Moustafa & Maida, 2007).

Activation levels of all units in the model are computed as follows:

$$A_j(t) = f\left(\sum_{i=1}^{n} u_{ji}(t) x_i(t)\right)$$

where u_{ji} is the weight connecting unit i to unit j, n is the number of units in the input; the input units x_i take binary (0,1) values (for similar simulation details see Barto, 1995; Schultz, Dayan, & Montague, 1997; Suri & Schultz, 1999); t is time step, f is the logistic sigmoid function:

$$f(x) = \frac{1}{1 + e^{-x}}$$

Weights are updated at every time step. Weight update is different in the different modules (see below).

Hippocampal region

The hippocampal region is implemented similarly as in previous work (A. A. Moustafa, et al., 2009) with one important exception: whereas the older models were trial-level, meaning that the hippocampal-region network learned to predict the US based on current inputs, the new version simulates within-trial events. Specifically, a conditioning trial is conceived as a series of events occurring at specified time points; at the start of a trial, tonic contextual cues

are presented (6 time steps in simulations presented below); CSs may appear and remain on for some time (3 time steps in the current simulations), and a US may appear at the end of the CS presentation (one time step). Our model proposes that the hippocampal region plays a general role in learning to use context to disambiguate competing responses to a stimulus.

The hippocampal module is a single layer of 20 nodes, and each node receives projections from the input layer that include nodes for sensory information, including CSs, US, and contexts. Learning in this network is Hebbian. The weight update rule here is as follows:

$$w_{ji}(t+1) = w_{ji}(t) + \alpha_{\text{hipp}} x_i(t) y_j(t)$$

where α_{hipp} is the learning rate for the hippocampal module (0.01 in the simulations presented here); x_i is the cortical input unit i; y_j is the activation of the unit j in the hippocampal layer. A k-winner-take-all network is used to simulate lateral inhibitory connections among nodes in the hippocampal module. As discussed above, hippocampal dynamics allows it to represents contextual information. The hippocampal representation (a vector composed of the values y_j) is projected to both BLA and vmPFC layers.

Amygdala and vmPFC

We simulate three subregions of the amygdala: BLA, ITC, and CeA. The simulated CeA is one node that integrates opposing excitatory (sensory-evoked BLA) and inhibitory (vmPFC-driven ITC) inputs:

$$CeA(t) = \sum_i \text{BLA\_activation}_i(t) - \text{ITC\_activation}(t)$$

The CeA output represents the strength of the conditioned fear response. All simulation results presented below reports activation level of the CeA node. It is presumed that this CeA output travels to motor and autonomic centers that drive the behavioral and physiological correlates of the fear response, although this is beyond the scope of the model.

The BLA module is a single layer of 40 nodes, and each node receives projections from the input layer encoding sensory information, including the CS and US. The BLA is trained using the TD algorithm (Barto, 1995; Barto, et al., 1983; Witten, 1977). Let $TD(t)$ be the temporal difference error signal at time t (also known as the effective reinforcement); $R(t)$ be the fear US presented at time t (it is 1 when fear US is presented and is 0 otherwise); $P(t)$ be the US prediction at time t; γ be the discount factor (which determines how expectation of future US affects current US predictions and which is set to 0.99 in all simulation runs presented here). The TD error is computed as follows:

$$TD(t) = R(t) + \gamma P(t) - P(t-1)$$

Let n be the number of input nodes (here, n=40); w_i be the weight from input node i; and x_i be activation of input units (which take binary (0,1) values). Prediction of fear US, $P(t)$, is computed as follows:

$$P(t) = \sum_{i=1}^{n} w_i(t) x_i(t)$$

Now, we describe the equations of the BLA. Let w_{ji} be the weight connecting unit i to BLA unit j:

$$w_{ji}(t+1) = w_{ji}(t) + \alpha\, TD(t)\, x_i(t)\, y_i(t)$$

where   is the learning rate; x_i is the cortical input unit i; y_j is the activation of the postsynaptic unit j in the BLA module.

For the simulated vmPFC, number of nodes and learning rules are similar to those described above for the BLA, except that learning in the vmPFC relies on negative TD errors (see equations above). We simulate ITC as a single node that receives fixed weights from the vmPFC.

Specifically, both the vmPFC and BLA integrate hippocampal and sensory input, and thus learn CS-US or CS-no US associations along with the contextual signal. As mentioned above, the main different between BLA and vmPFC is BLA is trained using positive TD error signals, while vmPFC is trained using negative TD errors signals.

Simulated lesions

We simulated lesions of a brain area by adding noise to activation levels of the simulated area, as previously done in other computational models (Joanisse & Seidenberg, 1999; Monaghan & Shillcock, 1998, 2008; Moustafa & Gluck, 2011; Olson & Humphreys, 1999; Ueno, Saito, Rogers, & Lambon Ralph, 2011). Adding noise to the output of a simulated brain region interferes with the output and damages the function of this region. Let   be Gaussian noise, A_j^{pre} be activation before adding noise to a brain area's activation levels; A_j^{post} be activation after adding noise to a brain area's activation. Then, for every unit j in the brain area, simulated lesion is conducted as follows:

$$A_j^{\text{post}} = A_j^{\text{pre}} + \delta_j$$

3-Results

We first present simulation results of the intact model. We show how the model simulates performance in fear conditioning, discrimination, extinction, reacquisition, and context shifting. We then present the effects of lesioning different parts of the model on fear acquisition and extinction.

Fear conditioning in intact animals

a- Discrimination learning—In rabbit heart rate conditioning, there is an increase in conditioned responding (bradycardic CR) to a stimulus CS that is paired with the US (i.e., CS+ = CS presented coterminous with US) but not to a stimulus that is not paired with the US (i.e., CS− = CS presented but no US) (Weidemann & Kehoe, 2003). The model similarly shows an increase in response to the CS+, but not to the CS− (Figure 3).

We have also measured CR to contextual cues during the pre-CS interval. In agreement with empirical results (Phillips & LeDoux, 1992), simulation results show that CR to the context is weaker than CR to the cue (see Figure 3, also see figure below). However, our model makes the novel prediction that the response to context will *decrease* with extended training (20 trials or more). We argue that with extensive training, animals learn to differentiate contexts from cues (i.e., they show stronger fear responses to the cue than to context).

Within a differential conditioning procedure, this would be expected because the CS+ is uniquely associated with the US whereas the context is not, i.e., the context is also associated with the CS− and absence of a US.

b- Acquisition and extinction—Once a CR has been acquired, further presentations of the CS alone (without US) result in gradual extinction of the CR (Leung, et al., 2007). Figure 4b shows that the model exhibits extinction, and that extinction is quicker when the context is changed between acquisition and extinction phases. This is in agreement with empirical data (Corcoran & Maren, 2001; Huff et al.; Orsini, Kim, Knapska, & Maren, 2011).

In addition, to test if extinction is CS-specific, we presented the model with intermixed trials with two stimuli (CS1+ and CS2+) in the same context. In the second phase, we extinguished CS1 only. In the third phase, we presented CS2 to test for the selectivity of extinction. Our results shows that the model correctly learns and then extinguishes a fear response to CS1 in phases 1 and 2 (see Figure 5). In the third phase, the model continues to show a fear response to CS2, following extinction of CS1, indicating stimulus-selectivity of learning and extinction in the model.

c- Reacquisition—Animals trained in one context, then extinguished in a new context, show faster relearning of the CR in the original training context, compared with animals retrained in the extinction context (Bouton & King, 1983; Bouton, et al., 2006). To simulate this effect, we trained the model to acquire a fear response in one context, followed by extinction trials in a different context. In a third, reacquisition phase, in some runs, the model was retrained in the acquisition context, and in other runs, the model was retrained in the extinction context. Simulation results show that that reacquisition is quicker in the acquisition context than in the extinction context (Figure 6). This demonstrates contextual sensitivity of learning and extinction in the model.

As similar to the reacquisition paradigm, our model also simulates performance in the fear renewal (ABA) paradigm, which is similar to the reacquisition paradigm, except that in the last phase the CS is presented without the US (see the Model Limitations section below on discussion of other renewal paradigms).

d- Activation patterns in simulated brain areas during acquisition and extinction—Figure 7 shows activation patterns in the simulated hippocampal region, vmPFC, and amygdala (particularly BLA) during acquisition and extinction. The model shows that some BLA nodes increase their activation during acquisition, but not extinction (Figure 7a,b). As predicted, the vmPFC shows the opposite pattern: vmPFC nodes increase their activation during extinction (Figure 7c,d). These simulation results are in agreement with studies showing BLA and vmPFC are involved in acquisition and extinction, respectively. Interestingly, the simulated hippocampal region shows different activation patterns during both acquisition and extinction (see Figure 8). Specifically, when the context is the same in the acquisition and extinction phases, hippocampal nodes that show increased activation in acquisition also showed increased activation in extinction, an indication of continued learning to represent contextual information (Figure 8a,b). However, when different contexts are used during acquisition and extinction, some hippocampal nodes show activation during acquisition but not during extinction, or vice versa (Figure 8c,d), although some hippocampal nodes showed activation in both phases. Our simulation results are also in line with empirical results showing that different subpopulations of hippocampal nodes are active during fear acquisition and extinction learning (Huh, et al., 2009; Tronson et al., 2009).

Lesion studies

Below, we present simulation results on the effects of lesioning BLA, hippocampal region, and vmPFC on fear conditioning and extinction.

a- BLA lesion—Simulation results show that lesioning the BLA interferes with fear acquisition, but not fear extinction (Figure 9). In the model, we obtain this result because lesioning the BLA affects CS-US association learning based on positive TD error sginals. This is in agreement with empirical results (Anglada-Figueroa & Quirk, 2005; LeDoux, 1993).

b- vmPFC lesion—Our simulation results show that vmPFC lesion (unlike BLA lesion) spares fear acquisition, but impairs fear extinction in later trials of the extinction phase (Figure 10). Interestingly, lesion studies have shown that extinction deficits appear in later trials of the extinction phase or extinction recall (Lebron, et al., 2004).

c- Hippocampal lesion—The effect of hippocampal lesion on fear conditioning is more complex. In agreement with empirical studies (Corcoran & Maren, 2001; Ji & Maren, 2007; Phillips & LeDoux, 1992), our simulation results show that lesioning the hippocampus does not interfere with simple cued fear acquisition and only mildly impairs extinction. However, lesioning the hippocampus significantly impairs extinction in a different context (Figure 11).

We obtain these results because both vmPFC and BLA circuits in the model are capable of learning CS-US and CS-no US associations without hippocampal input. However, once the context is important in some learning paradigm, lesioning the hippocampus leads to a behavioral effect. Specifically, in the intact model, extinction is faster in a new context, presumably because the new context generates a distinct representation of CS in context B in the hippocampal module that is different from CS in context A formed during acquisition, and thus it's easier to learn the new CS– no US association to the new CS-in-B representation. Given that the hippocampal module is important for the formation of contextual information, lesioning the hippocampus in our model interferes with extinction in a different context. Our simulation results are also in agreement with rabbit and rat data which show that hippocampal lesion or inactivation spares heart rate conditioning, but slows extinction (Chan, Jarrard, & Davidson, 2003; Powell & Buchanan, 1980).

4-Discussion

Our model accounts for various behavioral results on the role of amygdala, hippocampus, and vmPFC in fear conditioning and extinction. Our model is in agreement with data showing that the amygdala is a key structure mediating conditioned fear acquisition and expression (LeDoux, 1993), the hippocampus is essential for contextual processing, and vmPFC is important for forming fear extinction memories (for a review see Jovanovic & Ressler, 2010).

In addition, our model also accounts for other fear conditioning phenomena, including context shifting and reacquisition. It accounts for the data showing that extinction is not merely unlearning or forgetting of a CS-US association; rather it represents the acquisition of a competing CS-no US association (Bouton, 2004; Delamater, 2004). Our model also suggests that the reacquisition and renewal effect is obtained as a consequence of forming extinction memory. Forming extinction memory keeps acquisition memory intact, which explains accelerated re-emergence of the learned CR in the last phase of the reacquisition paradigm.

Our simulation results for context shift during extinction are in agreement with empirical findings in animals (Corcoran & Maren, 2001; Orsini, et al., 2011) and humans (Huff, et al., 2011). However, there are conflicting empirical results for the effects of context shifts during extinction (Delamater, 2004). This may be due to variability within some parameters such as the speed of learning or the number of trials in the first context before shifting to the new context. The same applies to contextual shifting during acquisition: while context shift may not impair learning (Bouton, 1993), some have found that context shift during learning leads to a decrease in responding in the new context (Penick & Solomon, 1991).

Our computational model proposes that fear responses (including freezing and heart rate changes) are driven by CeA outputs (for empirical support, see LeDoux, et al., 1988). This model feature is in agreement with studies showing that CeA is essential for driving freezing (Duvarci, et al., 2011) and heart rate (Kapp, et al., 1979) responses. In sum, we here argue that CeA neurons mediate the various fear responses, perhaps through interactions with different efferent hypothalamic and brainstem structures, each driving freezing and heart rate responses among others (for similar ideas, see Francis, Hernandez, & Powell, 1981; Hitchcock, Sananes, & Davis, 1989). For example, lesioning the periaqueductal gray area (which receive projections from the CeA) abolish freezing (LeDoux, et al., 1988), but not heart rate responses (Wilson & Kapp, 1994).

Few studies have found that lateral amygdala is involved in fear extinction learning (Quirk, Armony, & LeDoux, 1997). This is in contrast with recent lesion studies (Anglada-Figueroa & Quirk, 2005) as well as our simulation results. It is possible that early lesion studies of lateral amygdala interfered with ITC neurons, which lie at the border of the BLA. This could in turn impair extinction learning. Future studies should specifically investigate effects of BLA vs. ITC lesions on fear conditioning and extinction within a single paradigm.

In our model, activity of the hippocampus, BLA and vmPFC are impacted differently by presentations of the US. More specifically, US presentations in our model help increase the sparse representations of input stimuli in the hippocampal region, which is in agreement with existing neurophysiological studies reporting that hippocampal region neurons respond to the US presentation in conditioning tasks (but also see Barot, Chung, Kim, & Bernstein, 2009; McEchron & Disterhoft, 1997, 1999; Sugase-Miyamoto & Richmond, 2007). In other words, cues associated with the US will be represented differently from cues that are not associated with the US. This mechanism allows the model to efficiently represent different contexts using different sets of nodes (as in acquisition and extinction in different contexts). In contrast, US presentations will influence activity in BLA and vmPFC so as to support associative learning of the CS-US and CS-no US relationships.

Biological data supporting features of the model

Here, we discuss relevant biological studies that support features of our models. Our model proposes that two different neural pathways are responsible for fear acquisition and extinction. Early studies have shown that the BLA is critical for the acquisition of fear responses (Goldstein, 1965; Robinson, 1963). More recently, physiological and human brain imaging studies have shown that cortical structures, particularly vmPFC, also play a role in fear conditioning. Recent work by Pare and colleagues and Quirk colleagues suggests that vmPFC projection to GABAergic ITC may be responsible for fear inhibition and fear extinction (Amano, et al., 2010; Amir, et al., 2011; Quirk & Gehlert, 2003). Results from our computational model are consistent with these empirical studies, which suggest that there are two pathways for fear acquisition and extinction.

As shown above, we have trained the BLA and vmPFC areas using the TD model. Although it is true that most of the existing literature relates the TD model to reward-based learning,

the TD model is an error-based US-prediction learning model. It is very likely that many brain areas learn through predictions of subsequent events. In addition to our work, some prior empirical and modeling studies have related TD modeling to fear conditioning (Cole & McNally, 2009; McNally, et al., 2011; Spoormaker, Andrade, et al., 2011).

Comparison to other models

To our knowledge, our model is the first to simulate interactions among amygdala, vmPFC, ITC, and hippocampus in fear conditioning and extinction.

This model builds on our earlier models in which the hippocampus plays an essential role in representational and contextual processing (Gluck & Myers, 1993; A. A. Moustafa, et al., 2009). Our earlier models, however, did not simulate the role of the amygdala and vmPFC in conditioning. In addition, our earlier models were focused on simulating eyeblink conditioning data (A. A. Moustafa, Servatius, Pang, Gluck, & Myers, under review; Myers, Ermita, Hasselmo, & Gluck, 1998). Unlike eyeblink conditioning, fear conditioning is acquired in fewer trials (usually less than 10 or so trials). Accordingly, in model presented here, we have used a large learning rate than in previous models in order to simulate faster acquisition of fear responses as in empirical studies.

Below, we discuss models similar and related to our models, including (a) models of extinction and (b) model of amygdala function and fear conditioning.

a- Models of extinction—There are associative and nonassociative theories of extinction. Earlier associative theories of extinction argue that extinction involves erasure of memories, and weakening of the association between conditioned and unconditioned stimuli , which decrease the likelihood of producing a fear response (Mackintosh, 1975; Rescorla & Wagner, 1972). Unlike these theories, other theories and empirical data suggest that extinction training does not erase previously established memories, but rather lead to formation of a new inhibitory CS-no US memory (Bouton, 2002, 2004; Bouton & Bolles, 1979; Konorski, 1948; Pearce & Hall, 1980). As similar to our model, Suri and Schultz (1999) have also simulated extinction learning using the temporal difference model. However in their model, they have proposed that "extinction is erasure," and, thus, their model may not account for some phenomena including reacquisition, partial reinforcement extinction, or renewal effects. Unlike the Suri and Schultz model, we argue that positive and negative prediction errors impact different brain structures (for empirical evidence, see Spoormaker, Andrade, et al., 2011), and thus simulate data showing that fear extinction involves forming new memories that compete with fear memory.

Unlike the above-mentioned associative theories of extinction, nonassociative theories of extinction argue that extinction learning involves changes to attentional or stimulus competition processes (Gallistel & Gibbon, 2000; Kakade & Dayan, 2002; Redish, Jensen, Johnson, & Kurth-Nelson, 2007; Rescorla & Heth, 1975). For example, Pavlov (1927) argues that extinction training leads to a decrease in the activation of CS representation, and thus a decrease in fear response associated with it. Rescorla and Heth (1975) argue that extinction learning involves a decrease in US activation. An alternative model of extinction is proposed Ashby and colleagues (Ashby & Crossley, 2011a), in which extinction does not erase memories, but rather inhibit learning due to increased inhibition of interneurons in the striatum. Although the Ashby and Crossley model accounts for the reacquisition effect, it does not simulate extinction memory data put forth by physiologists (Milad & Quirk, 2002). A similar model, the O'Reilly and Munakata (2000) model, however, argues that weight values are not weakened by extinction trials, which allows their model to simulate the reacquisition effect since the first trial after extinction brings the synaptic strength above threshold and therefore reinstates the behavior.

Unlike previous models of extinction, which either simulate extinction as erasure of memories (for discussion see Suri & Schultz, 1999) or inhibition of learning (Ashby & Crossley; O'Reilly & Munakata, 2000), our model simulates the formation of extinction memory and shows how this mechanism can account for the reacquisition and contextual sensitivity of extinction.

b- Models of fear learning and amygdala—We note that although other computational models of the amygdala and fear conditioning exist, some have considered the role of cortical vs. thalamic inputs to amygdala (Armony, Servan-Schreiber, Cohen, & LeDoux, 1995, 1997; Armony, Servan-Schreiber, Romanski, Cohen, & LeDoux, 1997), and others have considered interactions between the amygdala and brain structures such as inferior temporal cortex (Rolls & Stringer, 2001), nucleus accumbens (Wagar & Thagard, 2003), and lateral hypothalamus (Dranias, Grossberg, & Bullock, 2008). These models did not specifically consider hippocampal-amygdala interactions.

A more recent model by Vlachos, Herry, Luthi, Aertsen, and Kumar (2011) simulates the role of the amygdala in fear conditioning and extinction. Like our model, this model proposed that different classes of neurons in the amygdala are important for both fear learning and extinction. Unlike our model, the Vlachos et al. model did not simulate the role of the hippocampus or vmPFC in fear conditioning. Recently, Li, Amano, Pare, and Nair (2011) have designed a biophysical model of ITC neurons' role in fear conditioning and extinction. The model also shows how increasing vmPFC (or infralimbic cortex) activity increases ITC activity, which in turn decreases fear responses. Unlike our model, this model did not simulate the role of hippocampus in contextual processes in fear conditioning; neither did this model explicitly simulate CeA function. Krasne et al.(2011) have also provided a model of the amygdala arguing that separate classes of BLA cells are important for fear extinction. As similar to the Li et al. (2011) model, the Krasne et al. model argues excitation and synaptic modification of interneurons in the BLA mediates extinction learning. This model does not simulate the role of vmPFC in fear extinction. Rudy and O'Reilly (1999, 2001) have presented models that focus on interactions among the hippocampus and cortex in contextual fear conditioning. Like our model, Rudy and O'Reilly argue that the hippocampus is important for conjunctive encoding and contextual fear conditioning. However, the Rudy and O'Reilly models did not simulate the role of vmPFC in fear extinction.

Unlike previous models of fear conditioning, our model (1) specifically considers the roles of hippocampus and amygdala subregions in fear conditioning, (2) examines their interactions with each other and with brain areas such as vmPFC that are also involved in associative learning, and (3) addresses the ways in which simulated lesions in one or more of these structures might give rise to the set of fear conditioning deficits observed in empirical studies.

Model limitations

One limitation of our model is that, as a systems-level model, it does not account for physiological details of the simulated brain areas. Similarly, the model has abstract nodes that do not simulate the biophysical properties of neurons (c.f. Li, et al., 2011). To our knowledge, prior biophysical models have mostly focused on the simulation of one brain region, such hippocampus (Rodriguez & Levy, 2001) or amygdala (Li, et al., 2011). However, there are pros and cons to simple models. Although simple models, including ours, may not capture some biological details (e.g., compare to Li, et al., 2011), they typically can explain performance in various behavioral tasks, which is the focus of our model. Specifically, the advantage of our simple model is that it allows us to focus on key

hypothesis regarding two pathways for positive and negative prediction errors using as few other free parameters as possible. We demonstrate that this simple model does, in fact, adequately simulate a large range of behavioral data. Nevertheless, future models should address how physiologically detailed models of the hippocampus, amygdala, and prefrontal interactions can simulate performance in a large number of behavioral studies, including conditioning data (as in our model).

Another limitation of the model is that it only provides a qualitative fit to existing fear conditioning and extinction data. This is because our model seeks to account for a broad range of experimental data that are related to lesion studies of different brain structures. Although lesion studies were conducted using different methods in empirical studies, our simulated lesion studies to all simulated brain structures—amygdala, hippocampal region, and vmPFC—were conducted in the same fashion and by using same noise values to induce lesion. This might explain the quantitative difference between our simulation results and existing empirical fear conditioning data. In short, Our model is intended to provide an abstract connectionist model that accommodates a variety of behavioral fear conditioning observations. In order to simulate a large set of behavioral and lesion data, we have used a simplified model that focuses more on the interactions among different brain regions rather than biophysical details of simulated brain areas.

Besides the hippocampus, amygdala, and vmPFC, other brain areas are also involved in fear conditioning and extinction, including the prelimbic cortex (Vidal-Gonzalez, et al., 2006) and dorsal anterior cingulate (Linnman, Rougemont-Bucking, Beucke, Zeffiro, & Milad, 2011). The anterior cingulate has been involved in conflict monitoring (Botvinick, Braver, Barch, Carter, & Cohen, 2001), and its role in fear conditioning could be related to solving a conflict a choice to activate either the CS-US or CS− no US pathway, so as to either initiate or inhibit a fear response. In addition, out model did not simulate the complex interactions among amygdala nuclei (for discussion see Manko, et al., 2011). Like most existing models (e.g., Krasne, et al., 2011; Vlachos, et al., 2011), representation of amygdala and hippocampus architecture in our model is simpler than known anatomy. For example, our model does not simulate the BLA-ITC connectivity (Sah, et al., 2003; Sotres-Bayon, Bush, & LeDoux, 2004). To our knowledge, not much is known about the function of this connectivity in fear conditioning or any other behavioral processes. Consequently, modeling such connectivity would be based entirely on speculation. Nonetheless, our model architecture is largely in agreement with previous theoretical models on the role of different amygdala nuclei in fear acquisition and extinction (see for example, Kim & Jung, 2006; Maren & Quirk, 2004).

Along the same lines, unlike eyeblink (Christian & Thompson, 2003) and reward-based instrumental conditioning (Schultz, et al., 1997), few studies have investigated the role of prediction error in fear learning (for example, see Spoormaker, Schroter, et al., 2011) and other studies testing the effects of negative and positive prediction errors on afferent neural structures, including hippocampus (Huh, et al., 2009; Tronson, et al., 2009). Various studies suggest that a complex neural network mediates the computation of prediction errors in fear learning, including periaqueductal gray and dopamine (McNally, et al., 2011), amygdala (Cole & McNally, 2009) and prefrontal cortex (Furlong, Cole, Hamlin, & McNally, 2010). Our model did not simulate the complex interactions among these brain structures in the computation of prediction errors in fear learning.

As mentioned above, theories that explain fear conditioning as both erasure of fear memories or forming extinction (inhibitory CS-no US) memories have some truth (for discussion see Pan, Schmidt, Wickens, & Hyland, 2008), depending on circumstances and experimental design. For example, studies have shown that fear extinction can lead to an

erasure of fear memories, depending on paradigm used (Schiller, Levy, Niv, LeDoux, & Phelps, 2008), amount of extinction training (Leung, et al., 2007), pharmacological usage (Han et al., 2009), or age of experimental rats (Quirk et al., 2010). Although here we have simulated fear extinction as forming inhibitory CS-no US memories, future modeling research should simulate the circumstances under which fear extinction can involve either erasure of fear memories or forming new inhibitory extinction memories.

Another limitation of our model is the finding that with more intense fear US, rats show stronger freezing to context than with a weaker foot shock (Phillips & LeDoux, 1992). Our model does not simulate the effects of US intensity on CR to context. Future work could expand the model to include the brainstem and other substrates where US intensity is processed and projected to the brain structures simulated here.

Model predictions

Despite the limitations, our model provides an initial framework for addressing interactions among the amygdala, hippocampus, and vmPFC in fear conditioning and extinction in animals.

In addition, as shown in Figure 5, our model predicts that animals can learn to differentiate safe from fear-eliciting responses. Specifically, our model predicts that if animals are trained with intermixed trials with two stimuli (CS1+ and CS2+), but subsequently are trained on CS1 extinction, in the third phase, associating CS2 with US will be learned quickly with no or minor interference from phase 2 extinction learning.

In addition, unlike some existing models (e.g., Gershman, Blei, & Niv, 2010), our model accounts for the partial reinforcement effect, whereby extinction is faster for fully, compared to partially, reinforced conditioning cues (Barad, 2006; Capaldi, 1958; Rescorla & Wagner, 1972). As suggested by Rescorla and Wagner (1972), our model also argues that the partial reinforcement effect is obtained as a result of forming inhibitory CS-no US memories. We are aware of only a few empirical studies that have used the partial reinforcement paradigm in fear conditioning (e.g., Milad et al., 2009; Milad et al., 2007). Importantly, our model shows the partial reinforcement effect only when using extensive training trials in the acquisition phase. This can perhaps explain why some empirical studies do not obtain the partial reinforcement effect. It is possible that the number of acquisition training trials influences whether or not the partial reinforcement effect is observed.

Our model proposes that negative prediction error activates the vmPFC. This model feature is supported by empirical evidence (see Cole & McNally, 2009; McNally, et al., 2011) and is also in agreement with empirical results showing that negative prediction errors activate medial prefrontal cortex areas (Brown & Braver, 2005; Holroyd & Coles, 2002). Our model additionally predicts that larger negative prediction errors in fear conditioning should be associated with greater vmPFC activation.

In sum, our model shows how interactions among the amygdala, vmPFC, and hippocampus explain various fear conditioning data, including fear acquisition and extinction, reacquisition, and contextual modulation of fear acquisition and extinction. Our model also accounts for data on how damage to the amygdala, vmPFC, or hippocampus interferes with fear conditioning and extinction.

Acknowledgments

5-References

Aggleton, JP. The Amygdala. Oxford University Press; Oxford: 2000.

Amano T, Unal CT, Pare D. Synaptic correlates of fear extinction in the amygdala. Nat Neurosci. 2010; 13(4):489–494. [PubMed: 20208529]

Amir A, Amano T, Pare D. Physiological Identification and Infralimbic Responsiveness of Rat Intercalated Amygdala Neurons. J Neurophysiol. 2011

Anagnostaras SG, Maren S, Fanselow MS. Temporally graded retrograde amnesia of contextual fear after hippocampal damage in rats: within-subjects examination. J Neurosci. 1999; 19(3):1106–1114. [PubMed: 9920672]

Anglada-Figueroa D, Quirk GJ. Lesions of the basal amygdala block expression of conditioned fear but not extinction. J Neurosci. 2005; 25(42):9680–9685. [PubMed: 16237172]

Armony JL, Servan-Schreiber D, Cohen JD, LeDoux JE. An anatomically constrained neural network model of fear conditioning. Behav Neurosci. 1995; 109(2):246–257. [PubMed: 7619315]

Armony JL, Servan-Schreiber D, Cohen JD, Ledoux JE. Computational modeling of emotion: explorations through the anatomy and physiology of fear conditioning. Trends Cogn Sci. 1997; 1(1):28–34. [PubMed: 21223850]

Armony JL, Servan-Schreiber D, Romanski LM, Cohen JD, LeDoux JE. Stimulus generalization of fear responses: effects of auditory cortex lesions in a computational model and in rats. Cereb Cortex. 1997; 7(2):157–165. [PubMed: 9087823]

Ashby FG, Crossley MJ. A computational model of how cholinergic interneurons protect striatal-dependent learning. J Cogn Neurosci. 2011a; 23(6):1549–1566. [PubMed: 20521851]

Ashby FG, Crossley MJ. A Computational Model of How Cholinergic Interneurons Protect Striatal-dependent Learning. J Cogn Neurosci. 2011b

Barad M. Is extinction of fear erasure or inhibition? Why both, of course. Learn Mem. 2006; 13(2): 108–109. [PubMed: 16585787]

Barot SK, Chung A, Kim JJ, Bernstein IL. Functional imaging of stimulus convergence in amygdalar neurons during Pavlovian fear conditioning. PLoS One. 2009; 4(7):e6156. [PubMed: 19582153]

Barto, AG. Adaptive critics and the basal ganglia. In: Houk, JC.; Davis, JL.; Beiser, DG., editors. Models of information processing in the basal ganglia. MIT Press; Cambridge, Mass.: 1995. p. xiip. 382

Barto AG, Sutton RS, Anderson CW. Neuronlike elements that can solve difficult learning control problems. IEEE Transactions on System, Man, and Cybernetics. 1983; 13:835–846.

Bellgowan PS, Helmstetter FJ. Neural systems for the expression of hypoalgesia during nonassociative fear. Behav Neurosci. 1996; 110(4):727–736. [PubMed: 8864264]

Bellman, RE. Dynamic programming. Princeton University Press; Princeton: 1957.

Berretta S, Pantazopoulos H, Caldera M, Pantazopoulos P, Pare D. Infralimbic cortex activation increases c-Fos expression in intercalated neurons of the amygdala. Neuroscience. 2005; 132(4): 943–953. [PubMed: 15857700]

Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD. Conflict monitoring and cognitive control. Psychol Rev. 2001; 108(3):624–652. [PubMed: 11488380]

Bouton ME. Differential control by context in the inflation and reinstatement paradigms. Journal of Experimental Psychology: Animal Behavior Processes. 1984; 10(1):56–74.

Bouton ME. Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. Psychol Bull. 1993; 114(1):80–99. [PubMed: 8346330]

Bouton ME. Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. Biol Psychiatry. 2002; 52(10):976–986. [PubMed: 12437938]

Bouton ME. Context and behavioral processes in extinction. Learn Mem. 2004; 11(5):485–494. [PubMed: 15466298]

Bouton ME, Bolles RC. Role of conditioned contextual stimuli in reinstatement of extinguished fear. J Exp Psychol Anim Behav Process. 1979; 5(4):368–378. [PubMed: 528893]

Bouton ME, King DA. Contextual control of the extinction of conditioned fear: tests for the associative value of the context. J Exp Psychol Anim Behav Process. 1983; 9(3):248–265. [PubMed: 6886630]

Bouton ME, Peck CA. Spontaneous recovery in cross-motivational transfer (counterconditioning). Anim Learn Behav. 1989; 20:313–321.

Bouton ME, Westbrook RF, Corcoran KA, Maren S. Contextual and temporal modulation of extinction: behavioral and biological mechanisms. Biol Psychiatry. 2006; 60(4):352–360. [PubMed: 16616731]

Brown JW, Braver TS. Learned predictions of error likelihood in the anterior cingulate cortex. Science. 2005; 307(5712):1118–1121. [PubMed: 15718473]

Buchanan SL, Thompson RH, Maxwell BL, Powell DA. Efferent connections of the medial prefrontal cortex in the rabbit. Exp Brain Res. 1994; 100(3):469–483. [PubMed: 7529194]

Burgos-Robles A, Vidal-Gonzalez I, Santini E, Quirk GJ. Consolidation of fear extinction requires NMDA receptor-dependent bursting in the ventromedial prefrontal cortex. Neuron. 2007; 53(6): 871–880. [PubMed: 17359921]

Capaldi EJ. The effect of different amounts of training on the resistance to extinction of different patterns of partially reinforced responses. J Comp Physiol Psychol. 1958; 51(3):367–371. [PubMed: 13563711]

Cavada C, Company T, Tejedor J, Cruz-Rizzolo RJ, Reinoso-Suarez F. The anatomical connections of the macaque monkey orbitofrontal cortex. A review. Cereb Cortex. 2000; 10(3):220–242.

Chan KH, Jarrard LE, Davidson TL. The effects of selective ibotenate lesions of the hippocampus on conditioned inhibition and extinction. Cogn Affect Behav Neurosci. 2003; 3(2):111–119. [PubMed: 12943326]

Christian KM, Thompson RF. Neural substrates of eyeblink conditioning: acquisition and retention. Learn Mem. 2003; 10(6):427–455. [PubMed: 14657256]

Cole S, McNally GP. Complementary roles for amygdala and periaqueductal gray in temporal-difference fear learning. Learn Mem. 2009; 16(1):1–7. [PubMed: 19117910]

Corbit LH, Balleine BW. Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. J Neurosci. 2005; 25(4): 962–970. [PubMed: 15673677]

Corcoran KA, Maren S. Hippocampal inactivation disrupts contextual retrieval of fear memory after extinction. J Neurosci. 2001; 21(5):1720–1726. [PubMed: 11222661]

Cox GE, Jordan D, Moruzzi P, Schwaber JS, Spyer KM, Turner SA. Amygdaloid influences on brain-stem neurones in the rabbit. J Physiol. 1986; 381:135–148. [PubMed: 3625533]

Davis M. The role of the amygdala in fear and anxiety. Annu Rev Neurosci. 1992; 15:353–375. [PubMed: 1575447]

Delamater AR. Experimental extinction in Pavlovian conditioning: behavioural and neuroscience perspectives. Q J Exp Psychol B. 2004; 57(2):97–132. [PubMed: 15204112]

Dranias MR, Grossberg S, Bullock D. Dopaminergic and non-dopaminergic value systems in conditioning and outcome-specific revaluation. Brain Res. 2008; 1238:239–287. [PubMed: 18674518]

Duvarci S, Popa D, Pare D. Central amygdala activity during fear conditioning. J Neurosci. 2011; 31(1):289–294. [PubMed: 21209214]

Francis J, Hernandez LL, Powell DA. Lateral hypothalamic lesions: effects on Pavlovian cardiac and eyeblink conditioning in the rabbit. Brain Res Bull. 1981; 6(2):155–163. [PubMed: 6970611]

Furlong TM, Cole S, Hamlin AS, McNally GP. The role of prefrontal cortex in predictive fear learning. Behav Neurosci. 2010; 124(5):574–586. [PubMed: 20939658]

Gallistel CR, Gibbon J. Time, rate, and conditioning. Psychol Rev. 2000; 107(2):289–344. [PubMed: 10789198]

Gershman SJ, Blei DM, Niv Y. Context, learning, and extinction. Psychol Rev. 2010; 117(1):197–209. [PubMed: 20063968]

Gibbs CM, Prescott LB, Powell DA. A comparison of multiple-unit activity in the medial prefrontal and agranular insular cortices during Pavlovian heart rate conditioning in rabbits. Exp Brain Res. 1992; 89(3):599–610. [PubMed: 1644124]

Gluck MA, Myers CE. Hippocampal mediation of stimulus representation: a computational theory. Hippocampus. 1993; 3(4):491–516. [PubMed: 8269040]

Goldstein ML. Effects of Hippocampal, Amygdala, Hypothalamic and Parietal Lesions on a Classically Conditioned Fear Response. Psychol Rep. 1965; 16:211–219. [PubMed: 14283965]

Han JH, Kushner SA, Yiu AP, Hsiang HL, Buch T, Waisman A, et al. Selective erasure of a fear memory. Science. 2009; 323(5920):1492–1496. [PubMed: 19286560]

Herry C, Ciocchi S, Senn V, Demmou L, Muller C, Luthi A. Switching on and off fear by distinct neuronal circuits. Nature. 2008; 454(7204):600–606. [PubMed: 18615015]

Hitchcock JM, Sananes CB, Davis M. Sensitization of the startle reflex by footshock: blockade by lesions of the central nucleus of the amygdala or its efferent pathway to the brainstem. Behav Neurosci. 1989; 103(3):509–518. [PubMed: 2544201]

Hobin JA, Ji J, Maren S. Ventral hippocampal muscimol disrupts context-specific fear memory retrieval after extinction in rats. Hippocampus. 2006; 16(2):174–182. [PubMed: 16358312]

Holroyd CB, Coles MG. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychol Rev. 2002; 109(4):679–709. [PubMed: 12374324]

Huff NC, Hernandez JA, Fecteau ME, Zielinski DJ, Brady R, Labar KS. Revealing context-specific conditioned fear memories with full immersion virtual reality. Front Behav Neurosci. 2011; 5:75. [PubMed: 22069384]

Huh KH, Guzman YF, Tronson NC, Guedea AL, Gao C, Radulovic J. Hippocampal Erk mechanisms linking prediction error to fear extinction: roles of shock expectancy and contextual aversive valence. Learn Mem. 2009; 16(4):273–278. [PubMed: 19318469]

Indovina I, Robbins TW, Nunez-Elizalde AO, Dunn BD, Bishop SJ. Fear-conditioning mechanisms associated with trait vulnerability to anxiety in humans. Neuron. 2011; 69(3):563–571. [PubMed: 21315265]

Jay TM, Witter MP. Distribution of hippocampal CA1 and subicular efferents in the prefrontal cortex of the rat studied by means of anterograde transport of Phaseolus vulgaris-leucoagglutinin. J Comp Neurol. 1991; 313(4):574–586. [PubMed: 1783682]

Ji J, Maren S. Electrolytic lesions of the dorsal hippocampus disrupt renewal of conditional fear after extinction. Learn Mem. 2005; 12(3):270–276. [PubMed: 15930505]

Ji J, Maren S. Hippocampal involvement in contextual modulation of fear extinction. Hippocampus. 2007; 17(9):749–758. [PubMed: 17604353]

Joanisse MF, Seidenberg MS. Impairments in verb morphology after brain injury: a connectionist model. Proc Natl Acad Sci U S A. 1999; 96(13):7592–7597. [PubMed: 10377460]

Jovanovic T, Ressler KJ. How the neurocircuitry and genetics of fear inhibition may inform our understanding of PTSD. Am J Psychiatry. 2010; 167(6):648–662. [PubMed: 20231322]

Kakade S, Dayan P. Acquisition and extinction in autoshaping. Psychol Rev. 2002; 109(3):533–544. [PubMed: 12088244]

Kapp BS, Frysinger RC, Gallagher M, Haselton JR. Amygdala central nucleus lesions: effect on heart rate conditioning in the rabbit. Physiol Behav. 1979; 23(6):1109–1117. [PubMed: 542522]

Kentros CG, Agnihotri NT, Streater S, Hawkins RD, Kandel ER. Increased attention to spatial context increases both place field stability and spatial memory. Neuron. 2004; 42(2):283–295. [PubMed: 15091343]

Kim, Fanselow MS. Modality-specific retrograde amnesia of fear. Science. 1992; 256(5057):675–677. [PubMed: 1585183]

Kim, Jung MW. Neural circuits and mechanisms involved in Pavlovian fear conditioning: a critical review. Neurosci Biobehav Rev. 2006; 30(2):188–202. [PubMed: 16120461]

Knierim JJ. Dynamic interactions between local surface cues, distal landmarks, and intrinsic circuitry in hippocampal place cells. J Neurosci. 2002; 22(14):6254–6264. [PubMed: 12122084]

Konorski, J. Conditioned reflexes and neuron organization. Cambridge University Press; New York: 1948.

Krasne FB, Fanselow MS, Zelikowsky M. Design of a neurally plausible model of fear learning. Front Behav Neurosci. 2011; 5:41. [PubMed: 21845175]

LaBar KS, Gatenby JC, Gore JC, LeDoux JE, Phelps EA. Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fMRI study. Neuron. 1998; 20(5):937–945. [PubMed: 9620698]

Lanuza E, Nader K, Ledoux JE. Unconditioned stimulus pathways to the amygdala: effects of posterior thalamic and cortical lesions on fear conditioning. Neuroscience. 2004; 125(2):305–315. [PubMed: 15062974]

Lebron K, Milad MR, Quirk GJ. Delayed recall of fear extinction in rats with lesions of ventral medial prefrontal cortex. Learn Mem. 2004; 11(5):544–548. [PubMed: 15466306]

LeDoux JE. Brain mechanisms of emotion and emotional learning. Curr Opin Neurobiol. 1992; 2(2): 191–197. [PubMed: 1638153]

LeDoux JE. Emotional memory systems in the brain. Behav Brain Res. 1993; 58(1-2):69–79. [PubMed: 8136051]

LeDoux JE, Iwata J, Cicchetti P, Reis DJ. Different projections of the central amygdaloid nucleus mediate autonomic and behavioral correlates of conditioned fear. J Neurosci. 1988; 8(7):2517–2529. [PubMed: 2854842]

Leung HT, Bailey GK, Laurent V, Westbrook RF. Rapid reacquisition of fear to a completely extinguished context is replaced by transient impairment with additional extinction training. J Exp Psychol Anim Behav Process. 2007; 33(3):299–313. [PubMed: 17620028]

Leutgeb JK, Leutgeb S, Moser MB, Moser EI. Pattern separation in the dentate gyrus and CA3 of the hippocampus. Science. 2007; 315(5814):961–966. [PubMed: 17303747]

Li G, Amano T, Pare D, Nair SS. Impact of infralimbic inputs on intercalated amygdala neurons: a biophysical modeling study. Learn Mem. 2011; 18(4):226–240. [PubMed: 21436395]

Likhtik E, Pelletier JG, Paz R, Pare D. Prefrontal control of the amygdala. J Neurosci. 2005; 25(32): 7429–7437. [PubMed: 16093394]

Likhtik E, Popa D, Apergis-Schoute J, Fidacaro GA, Pare D. Amygdala intercalated neurons are required for expression of fear extinction. Nature. 2008; 454(7204):642–645. [PubMed: 18615014]

Linnman C, Rougemont-Bucking A, Beucke JC, Zeffiro TA, Milad MR. Unconditioned responses and functional fear networks in human classical conditioning. Behav Brain Res. 2011; 221(1):237–245. [PubMed: 21377494]

Mackintosh NJ. A theory of attention: Variations in the associability of stimuli with reinforcement. Psychological Review. 1975; 82:276–298.

Manko M, Geracitano R, Capogna M. Functional connectivity of the main intercalated nucleus of the mouse amygdala. J Physiol. 2011; 589(Pt 8):1911–1925. [PubMed: 21224220]

Maren S. Overtraining does not mitigate contextual fear conditioning deficits produced by neurotoxic lesions of the basolateral amygdala. J Neurosci. 1998; 18(8):3088–3097. [PubMed: 9526025]

Maren S, Aharonov G, Fanselow MS. Retrograde abolition of conditional fear after excitotoxic lesions in the basolateral amygdala of rats: absence of a temporal gradient. Behav Neurosci. 1996; 110(4): 718–726. [PubMed: 8864263]

Maren S, Fanselow MS. Electrolytic lesions of the fimbria/fornix, dorsal hippocampus, or entorhinal cortex produce anterograde deficits in contextual fear conditioning in rats. Neurobiol Learn Mem. 1997; 67(2):142–149. [PubMed: 9075242]

Maren S, Quirk GJ. Neuronal signalling of fear memory. Nat Rev Neurosci. 2004; 5(11):844–852. [PubMed: 15496862]

McCabe PM, Gentile CG, Markgraf CG, Teich AH, Schneiderman N. Ibotenic acid lesions in the amygdaloid central nucleus but not in the lateral subthalamic area prevent the acquisition of differential Pavlovian conditioning of bradycardia in rabbits. Brain Res. 1992; 580(1-2):155–163. [PubMed: 1504795]

McCabe PM, McEchron MD, Green EJ, Schneiderman N. Electrolytic and ibotenic acid lesions of the medial subnucleus of the medial geniculate prevent the acquisition of classically conditioned heart

rate to a single acoustic stimulus in rabbits. Brain Res. 1993; 619(1-2):291–298. [PubMed: 8374784]

McEchron MD, Disterhoft JF. Sequence of single neuron changes in CA1 hippocampus of rabbits during acquisition of trace eyeblink conditioned responses. J Neurophysiol. 1997; 78(2):1030–1044. [PubMed: 9307133]

McEchron MD, Disterhoft JF. Hippocampal encoding of non-spatial trace conditioning. Hippocampus. 1999; 9(4):385–396. [PubMed: 10495020]

McNally GP, Johansen JP, Blair HT. Placing prediction into the fear circuit. Trends Neurosci. 2011; 34(6):283–292. [PubMed: 21549434]

Milad MR, Orr SP, Pitman RK, Rauch SL. Context modulation of memory for fear extinction in humans. Psychophysiology. 2005; 42(4):456–464. [PubMed: 16008774]

Milad MR, Pitman RK, Ellis CB, Gold AL, Shin LM, Lasko NB, et al. Neurobiological basis of failure to recall extinction memory in posttraumatic stress disorder. Biol Psychiatry. 2009; 66(12):1075–1082. [PubMed: 19748076]

Milad MR, Quirk GJ. Neurons in medial prefrontal cortex signal memory for fear extinction. Nature. 2002; 420(6911):70–74. [PubMed: 12422216]

Milad MR, Vidal-Gonzalez I, Quirk GJ. Electrical stimulation of medial prefrontal cortex reduces conditioned fear in a temporally specific manner. Behav Neurosci. 2004; 118(2):389–394. [PubMed: 15113265]

Milad MR, Wright CI, Orr SP, Pitman RK, Quirk GJ, Rauch SL. Recall of fear extinction in humans activates the ventromedial prefrontal cortex and hippocampus in concert. Biol Psychiatry. 2007; 62(5):446–454. [PubMed: 17217927]

Monaghan P, Shillcock R. The cross-over effect in unilateral neglect. Modelling detailed data in the line-bisection task. Brain. 1998; 121(Pt 5):907–921. [PubMed: 9619193]

Monaghan P, Shillcock R. Hemispheric dissociation and dyslexia in a computational model of reading. Brain Lang. 2008; 107(3):185–193. [PubMed: 18215416]

Montague PR, Dayan P, Sejnowski TJ. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J Neurosci. 1996; 16(5):1936–1947. [PubMed: 8774460]

Morgan MA, Romanski LM, LeDoux JE. Extinction of emotional learning: contribution of medial prefrontal cortex. Neurosci Lett. 1993; 163(1):109–113. [PubMed: 8295722]

Moustafa, Gluck. A neurocomputational model of dopamine and prefrontal-striatal interactions during multicue category learning by Parkinson patients. J Cogn Neurosci. 2011; 23(1):151–167. [PubMed: 20044893]

Moustafa, Maida AS. Using TD learning to simulate working memory performance in a model of the prefrontal cortex and basal ganglia. Cognitive Systems Research. 2007; 8:262–281.

Moustafa AA, Myers CE, Gluck MA. A neurocomputational model of classical conditioning phenomena: a putative role for the hippocampal region in associative learning. Brain Res. 2009; 1276:180–195. [PubMed: 19379717]

Moustafa AA, Servatius RJ, Pang KC, Gluck MA, Myers C,E. Why trace and delay eyeblink conditioning are sometimes (but not always) hippocampal dependent: A computational model. Psych Rev. (under review).

Moutoussis M, Bentall RP, Williams J, Dayan P. A temporal difference account of avoidance learning. Network. 2008; 19(2):137–160. [PubMed: 18569725]

Myers, Davis M. Mechanisms of fear extinction. Mol Psychiatry. 2007; 12(2):120–150. [PubMed: 17160066]

Myers, Ermita B, Hasselmo M, Gluck MA. Further implications of a computational model of septohippocampal cholinergic modulation in eyeblink conditioning. Psychobiology. 1998; 26(1):1–20.

O'Reilly, RC.; Munakata, Y. Computational explorations in cognitive neuroscience : understanding the mind by simulating the brain. MIT Press; Cambridge, Mass.: 2000.

Olson AC, Humphreys GW. Connectionist models of neuropsychological disorders. Trends Cogn Sci. 1999; 1(6):222–228. [PubMed: 21223911]

Orsini CA, Kim JH, Knapska E, Maren S. Hippocampal and prefrontal projections to the basal amygdala mediate contextual regulation of fear after extinction. J Neurosci. 2011; 31(47):17269–17277. [PubMed: 22114293]

Pan WX, Schmidt R, Wickens JR, Hyland BI. Tripartite mechanism of extinction suggested by dopamine neuron activity and temporal difference model. J Neurosci. 2008; 28(39):9619–9631. [PubMed: 18815248]

Pare D, Quirk GJ, Ledoux JE. New vistas on amygdala networks in conditioned fear. J Neurophysiol. 2004; 92(1):1–9. [PubMed: 15212433]

Pare D, Smith Y. The intercalated cell masses project to the central and medial nuclei of the amygdala in cats. Neuroscience. 1993; 57(4):1077–1090. [PubMed: 8309544]

Parkes SL, Westbrook RF. The basolateral amygdala is critical for the acquisition and extinction of associations between a neutral stimulus and a learned danger signal but not between two neutral stimuli. J Neurosci. 2010; 30(38):12608–12618. [PubMed: 20861367]

Pascoe JP, Kapp BS. Electrophysiological characteristics of amygdaloid central nucleus neurons during Pavlovian fear conditioning in the rabbit. Behav Brain Res. 1985; 16(2-3):117–133. [PubMed: 4041212]

Pavlov, IP. Conditioned reflexes. Oxford University Press; London: 1927.

Pearce JM, Hall G. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychol Rev. 1980; 87(6):532–552. [PubMed: 7443916]

Penick S, Solomon PR. Hippocampus, context, and conditioning. Behav Neurosci. 1991; 105(5):611–617. [PubMed: 1815614]

Phelps EA, Delgado MR, Nearing KI, LeDoux JE. Extinction learning in humans: role of the amygdala and vmPFC. Neuron. 2004; 43(6):897–905. [PubMed: 15363399]

Phillips RG, LeDoux JE. Differential contribution of amygdala and hippocampus to cued and contextual fear conditioning. Behav Neurosci. 1992; 106(2):274–285. [PubMed: 1590953]

Powell D, Buchanan S. Autonomic-somatic relationships in the rabbit (Oryctolagus cuniculus): Effects of hippocampal lesions. Physiological Psychology. 1980; 8:455–462.

Quirk GJ, Armony JL, LeDoux JE. Fear conditioning enhances different temporal components of tone-evoked spike trains in auditory cortex and lateral amygdala. Neuron. 1997; 19(3):613–624. [PubMed: 9331352]

Quirk GJ, Gehlert DR. Inhibition of the amygdala: key to pathological states? Ann N Y Acad Sci. 2003; 985:263–272. [PubMed: 12724164]

Quirk GJ, Likhtik E, Pelletier JG, Pare D. Stimulation of medial prefrontal cortex decreases the responsiveness of central amygdala output neurons. J Neurosci. 2003; 23(25):8800–8807. [PubMed: 14507980]

Quirk GJ, Mueller D. Neural mechanisms of extinction learning and retrieval. Neuropsychopharmacology. 2008; 33(1):56–72. [PubMed: 17882236]

Quirk GJ, Pare D, Richardson R, Herry C, Monfils MH, Schiller D, et al. Erasing fear memories with extinction training. J Neurosci. 2010; 30(45):14993–14997. [PubMed: 21068303]

Quirk GJ, Russo GK, Barron JL, Lebron K. The role of ventromedial prefrontal cortex in the recovery of extinguished fear. J Neurosci. 2000; 20(16):6225–6231. [PubMed: 10934272]

Redish AD, Jensen S, Johnson A, Kurth-Nelson Z. Reconciling reinforcement learning models with behavioral extinction and renewal: implications for addiction, relapse, and problem gambling. Psychol Rev. 2007; 114(3):784–805. [PubMed: 17638506]

Rescorla, Heth CD. Reinstatement of fear to an extinguished conditioned stimulus. J Exp Psychol Anim Behav Process. 1975; 1(1):88–96. [PubMed: 1151290]

Rescorla; Wagner, AR. A theory of Pavlovian conditioning:variations in the effectiveness of reinforcement and non-reinforcement. In: Black, AH.; Prokasy, WF., editors. Classical conditioning II: current research and theory. Appleton-Century-Crofts; New York: 1972. p. 64-99.

Resstel LB, Joca SR, Correa FM, Guimaraes FS. Effects of reversible inactivation of the dorsal hippocampus on the behavioral and cardiovascular responses to an aversive conditioned context. Behav Pharmacol. 2008; 19(2):137–144. [PubMed: 18332678]

Robinson E. Effect of Amygdalectomy on Fear-Motivated Behavior in Rats. J Comp Physiol Psychol. 1963; 56:814–820. [PubMed: 14050166]

Rodriguez P, Levy WB. A model of hippocampal activity in trace conditioning: where's the trace? Behav Neurosci. 2001; 115(6):1224–1238. [PubMed: 11770054]

Rolls ET, Stringer SM. A model of the interaction between mood and memory. Network. 2001; 12(2): 89–109. [PubMed: 11405424]

Rorick-Kehn LM, Steinmetz JE. Amygdalar unit activity during three learning tasks: eyeblink classical conditioning, Pavlovian fear conditioning, and signaled avoidance conditioning. Behav Neurosci. 2005; 119(5):1254–1276. [PubMed: 16300433]

Rosenkranz JA, Grace AA. Dopamine-mediated modulation of odour-evoked amygdala potentials during pavlovian conditioning. Nature. 2002; 417(6886):282–287. [PubMed: 12015602]

Rudy JW, Huff NC, Matus-Amat P. Understanding contextual fear conditioning: insights from a two-process model. Neurosci Biobehav Rev. 2004; 28(7):675–685. [PubMed: 15555677]

Rudy JW, O'Reilly RC. Contextual fear conditioning, conjunctive representations, pattern completion, and the hippocampus. Behav Neurosci. 1999; 113(5):867–880. [PubMed: 10571471]

Rudy JW, O'Reilly RC. Conjunctive representations, the hippocampus, and contextual fear conditioning. Cogn Affect Behav Neurosci. 2001; 1(1):66–82. [PubMed: 12467104]

Sah P, Faber ES, Lopez De Armentia M, Power J. The amygdaloid complex: anatomy and physiology. Physiol Rev. 2003; 83(3):803–834. [PubMed: 12843409]

Sananes CB, Campbell BA. Role of the central nucleus of the amygdala in olfactory heart rate conditioning. Behav Neurosci. 1989; 103(3):519–525. [PubMed: 2736066]

Santini E, Quirk GJ, Porter JT. Fear conditioning and extinction differentially modify the intrinsic excitability of infralimbic neurons. J Neurosci. 2008; 28(15):4028–4036. [PubMed: 18400902]

Schiller D, Levy I, Niv Y, LeDoux JE, Phelps EA. From fear to safety and back: reversal of fear in the human brain. J Neurosci. 2008; 28(45):11517–11525. [PubMed: 18987188]

Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. Science. 1997; 275(5306):1593–1599. [PubMed: 9054347]

Sierra-Mercado, Corcoran KA, Lebron-Milad K, Quirk GJ. Inactivation of the ventromedial prefrontal cortex reduces expression of conditioned fear and impairs subsequent recall of extinction. Eur J Neurosci. 2006; 24(6):1751–1758. [PubMed: 17004939]

Sierra-Mercado, Padilla-Coreano N, Quirk GJ. Dissociable roles of prelimbic and infralimbic cortices, ventral hippocampus, and basolateral amygdala in the expression and extinction of conditioned fear. Neuropsychopharmacology. 2010; 36(2):529–538. [PubMed: 20962768]

Sotres-Bayon F, Bush DE, LeDoux JE. Emotional perseveration: an update on prefrontal-amygdala interactions in fear extinction. Learn Mem. 2004; 11(5):525–535. [PubMed: 15466303]

Spoormaker VI, Andrade KC, Schroter MS, Sturm A, Goya-Maldonado R, Samann PG, et al. The neural correlates of negative prediction error signaling in human fear conditioning. Neuroimage. 2011; 54(3):2250–2256. [PubMed: 20869454]

Spoormaker VI, Schroter MS, Andrade KC, Dresler M, Kiem SA, Goya-Maldonado R, et al. Effects of rapid eye movement sleep deprivation on fear extinction recall and prediction error signaling. Hum Brain Mapp. 2011

Stefanacci L, Amaral DG. Some observations on cortical inputs to the macaque monkey amygdala: an anterograde tracing study. J Comp Neurol. 2002; 451(4):301–323. [PubMed: 12210126]

Sugase-Miyamoto Y, Richmond BJ. Cue and reward signals carried by monkey entorhinal cortex neurons during reward schedules. Exp Brain Res. 2007; 181(2):267–276. [PubMed: 17396249]

Suri RE, Schultz W. A neural network model with dopamine-like reinforcement signal that learns a spatial delayed response task. Neuroscience. 1999; 91(3):871–890. [PubMed: 10391468]

Sutton, RS.; Barto, AG. Reinforcement learning : an introduction. MIT Press; Cambridge, Mass.: 1998.

Swanson LW, Petrovich GD. What is the amygdala? Trends Neurosci. 1998; 21(8):323–331. [PubMed: 9720596]

Thorndike, EL. Animal intelligence. Hafner; Darien, CT: 1911.

Tronson NC, Schrick C, Guzman YF, Huh KH, Srivastava DP, Penzes P, et al. Segregated populations of hippocampal principal CA1 neurons mediating conditioning and extinction of contextual fear. J Neurosci. 2009; 29(11):3387–3394. [PubMed: 19295145]

Tronson NC, Wiseman SL, Olausson P, Taylor JR. Bidirectional behavioral plasticity of memory reconsolidation depends on amygdalar protein kinase A. Nat Neurosci. 2006; 9(2):167–169. [PubMed: 16415868]

Ueno T, Saito S, Rogers TT, Lambon Ralph MA. Lichtheim 2: synthesizing aphasia and the neural basis of language in a neurocomputational model of the dual dorsal-ventral language pathways. Neuron. 2011; 72(2):385–396. [PubMed: 22017995]

Vidal-Gonzalez I, Vidal-Gonzalez B, Rauch SL, Quirk GJ. Microstimulation reveals opposing influences of prelimbic and infralimbic cortex on the expression of conditioned fear. Learn Mem. 2006; 13(6):728–733. [PubMed: 17142302]

Vlachos I, Herry C, Luthi A, Aertsen A, Kumar A. Context-dependent encoding of fear and extinction memories in a large-scale network model of the Basal amygdala. PLoS Comput Biol. 2011; 7(3):e1001104. [PubMed: 21437238]

Wagar BM, Thagard P. Using computational neuroscience to investigate the neural correlates of cognitive-affective integration during covert decision making. Brain Cogn. 2003; 53(2):398–402. [PubMed: 14607190]

Weidemann G, Kehoe EJ. Savings in classical conditioning in the rabbit as a function of extended extinction. Learn Behav. 2003; 31(1):49–68. [PubMed: 18450069]

Weisz DJ, Harden DG, Xiang Z. Effects of amygdala lesions on reflex facilitation and conditioned response acquisition during nictitating membrane response conditioning in rabbit. Behav Neurosci. 1992; 106(2):262–273. [PubMed: 1317182]

Wiersma A, Bohus B, Koolhaas JM. Corticotropin-releasing hormone microinfusion in the central amygdala diminishes a cardiac parasympathetic outflow under stress-free conditions. Brain Res. 1993; 625(2):219–227. [PubMed: 8275304]

Wilson A, Kapp BS. Effect of lesions of the ventrolateral periaqueductal gray on the Pavlovian conditioned heart rate response in the rabbit. Behav Neural Biol. 1994; 62(1):73–76. [PubMed: 7945148]

Witten IH. An adaptive optimal controller for discrete-time Markov environments. Information and Control. 1977; 34:286–295.

Zelikowsky M, Pham DL, Fanselow MS. Temporal factors control hippocampal contributions to fear renewal after extinction. Hippocampus. 2011

Below, are the highlights of our study

- We present a systems-level model of fear conditioning.

- We simulate function of the amygdala, hippocampus, and prefrontal cortex.

- The model accounts fear conditioning data, including acquisition and extinction.

- Our model also assumes that fear extinction involves inhibitory CS-no US learning.

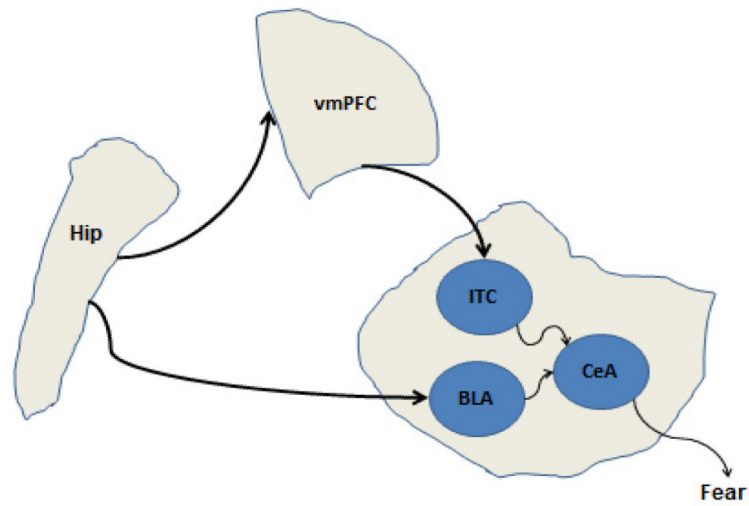- Our model assumes two subregions within the amygdala that separately subserve fear acquisition and extinction.

Figure 1.
A simplified anatomical diagram of connections among the hippocampus, vmPFC, amygdala, brain structures involved in fear conditioning. The hippocampal region (hip) projects to both vmPFC and amygdala. The amygdala has different nuclei; each has different connectivity patterns to afferent structures, and plays different function in fear conditioning (especially heart rate and freezing responses).
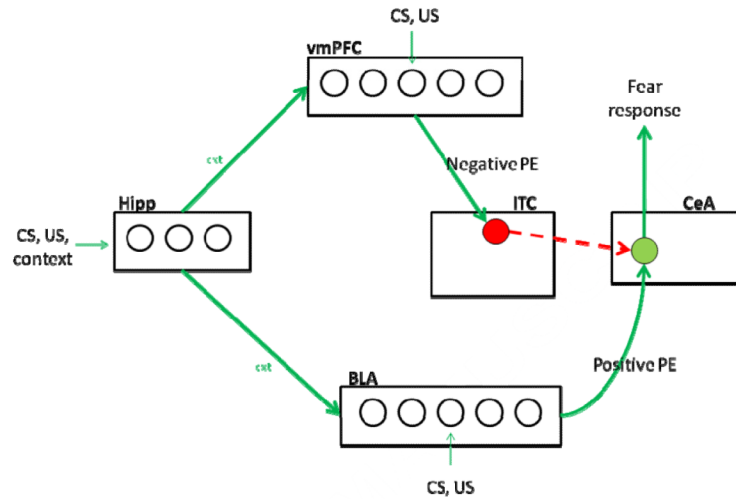
Figure 2.
Model architecture. The model simulates interactions among the amygdala, hippocampal region (HR), and ventromedial prefrontal cortex (vmPFC) in fear conditioning in rats. Our model simulates three subregions of the amygdala including basolateral amygdala (BLA), central nucleus (CEA), and intercalated cells (ITC). In our model, the BLA is essential for fear acquisition, while the ITC (along with vmPFC) is important for fear extinction. The CEA in the model initiates fear CRs. The hippocampal region (Hipp) learns to form representations of contextual information. The ventromedial prefrontal cortex (vmPFC) receives CS, US and hippocampal inputs, and outputs a "safety" signal which excites intercalated cells (ITC) in the amygdala that inhibit the CEA from driving fear CRs. The BLA and vmPFC are trained using the temporal difference (TD) model. Specifically, positive prediction (PE) errors participates in training the BLA to form conditioned fear CRs, while negative PE train the vmPFC to form safety signals by decreasing the probability of CEA to fire.
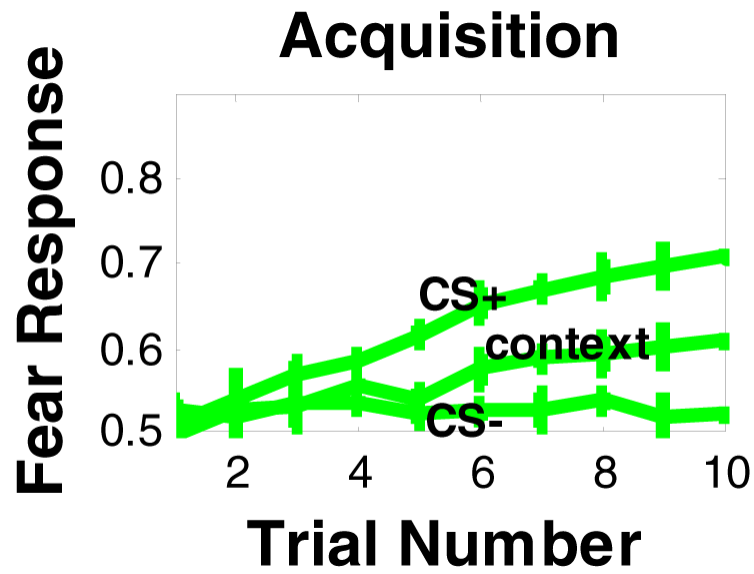
Figure 3.
Simulation results of fear conditioning and discrimination, using the model in Figure 1.The model shows increased fear response to CS associated with fear-inducing US (CS+), but not to CS−, that is not associated with fear US. The model also shows weaker response to context rather than to CS+.
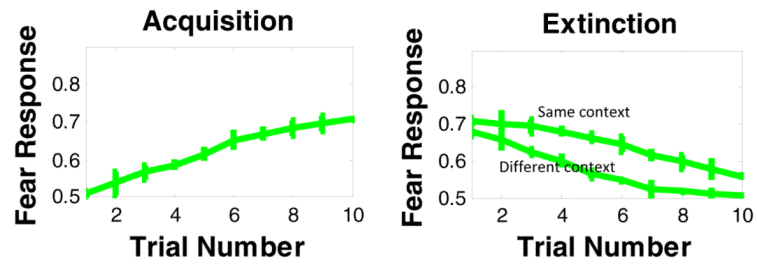
Figure 4.
Simulation results of fear conditioning and extinction in the same or different context. The model also accounts for fear extinction: simulations show that fear responses decreases when US is presented in the extinction phase. Our model shows that extinguishing fear responses is faster when context used in extinction trials is different from context used in acquisition.
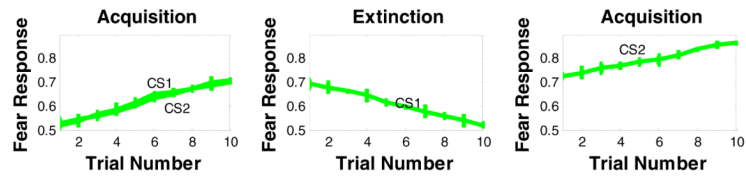
Figure 5.
Simulation results of the effects of training the model to acquire fear responses to two stimuli (CS1 & CS2), but then extinguishing fear responses to one cue only (CS1). Results show continued strong responding to CS2 following extinction of CS1, demonstrating stimulus-selectivity of acquisition and extinction in the model.
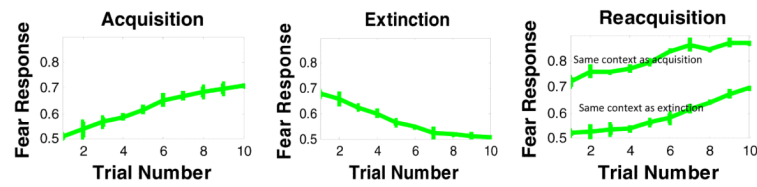
Figure 6.
Simulation results of the reacquisition effect. Our model shows that reacquisition of fear responses depends on whether acquisition or extinction context is used. Specifically, reacquisition of fear responses is faster when acquisition rather than extinction context is used.
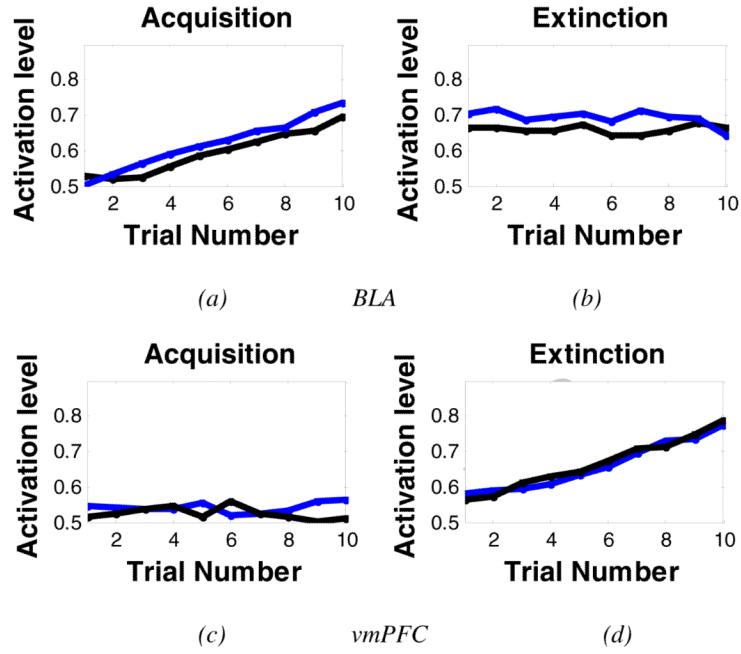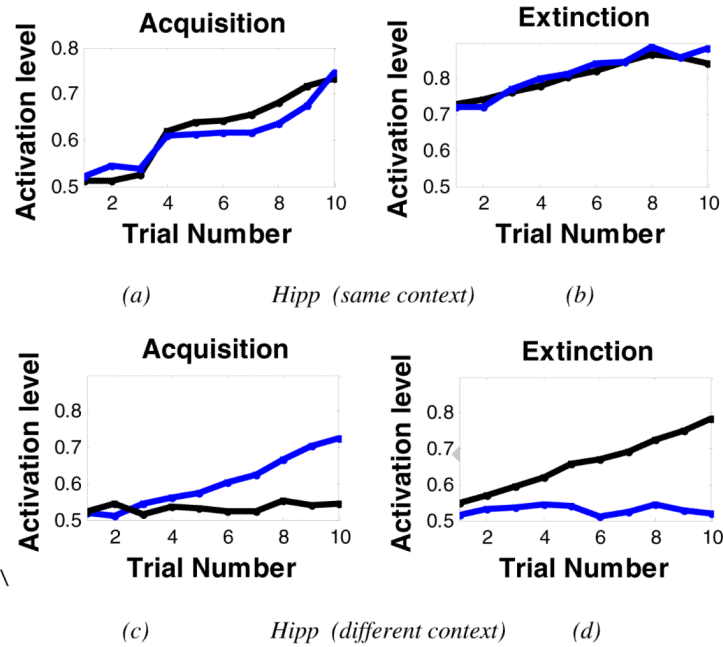
Figure 7.
Activation levels of the simulated vmPFC and amygdala during fear acquisition and extinction. Each color here represents a different node in the simulated brain area. (a,b) figures show activation from two representative BLA nodes during acquisition and extinction. BLA nodes showed increased activation during acquisition, but activation levels did not change during extinction. (c,d) figures show activation from two vmPFC nodes during acquisition and extinction. The vmPFC shows the opposite activation patterns: increased activation during extinction but not acquisition. Some other nodes (not shown) in the simulated brain areas did not show any significant change during acquisition or extinction.

Figure 8.
Activation levels of the simulated hippocampal region during fear acquisition and extinction, using same and different context in the extinction phase. As above, here, each color here represents a different node in the simulated brain area. (a,b) Using same context in fear acquisition and extinction. In the simulated hippocampal region, nodes that show increased activation in acquisition also showed increased activation in extinction, an indication of learning to represent contextual information. (c,d) Using different contexts in fear acquisition and extinction. When context is changed in the extinction phase, hippocampal nodes that were active during acquisition are not active in extinction (see blue node), but other hippocampal nodes become active during extinction (black node). This indicates that different subpopulations of hippocampal nodes encode different contexts in fear acquisition and extinction. This differential pattern of activation is not found in the simulated BLA or vmPFC modules (see previous figure).
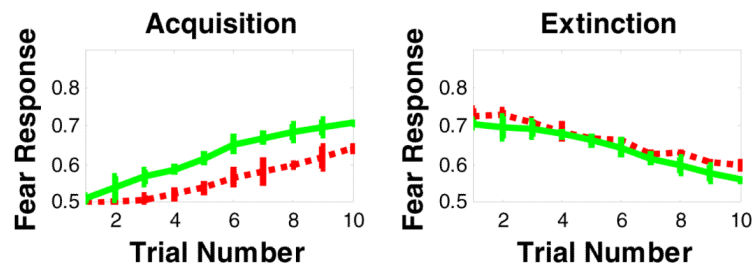
Figure 9.
Simulation results of the effects of lesioning the BLA on fear acquisition and extinction. Lesioning BLA in our model impairs fear acquisition, but does not affect fear extinction. The green curve refers to intact model, while the red curve refers to the BLA-lesioned model.
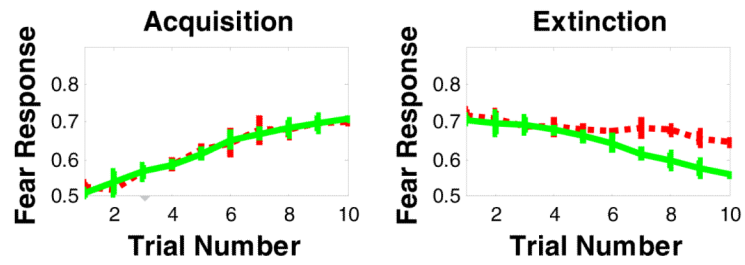
Figure 10.
Simulation results of the effects of lesioning the vmPFC on fear acquisition and extinction. Lesioning vmPFC in our model impairs fear extinction, but does not affect fear acquisition. The green curve refers to intact model, while the red curve refers to the vmPFC-lesioned model.
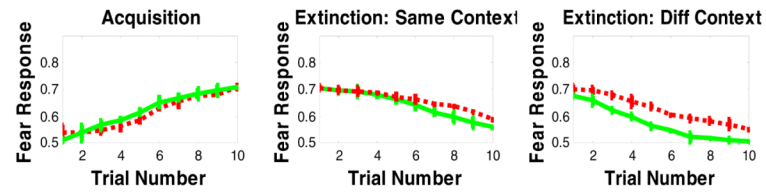
Figure 11.
Simulation results of the effects of lesioning the hippocampal region on fear acquisition and extinction. Lesioning the hippocampus in our model slightly impairs fear acquisition and extinction in same context as acquisition, but significantly impairs extinction in a different context. The green curve refers to intact model, while the red curve refers to the hippocampal-lesioned model.

Table 1

Tasks simulated in the model. In all experiments, contexts are referred to as A and B, cues as X. "AX-" means X is presented in context A with no US, while AX+ means X is presented in context A with the US in the same trial. ';' separates different trials in the same phase. In each of these phases, the corresponding context is presented to the network by itself before cue presentation mimicking the presence of animal inside a box (see Experimental Procedure).

Simulation	Phase 1	Phase 2	Phase 3
Fear conditioning and extinction	AX+	AX−	
Extinction in a new context	AX+	BX−	
Renewal	AX+	BX−	AX−; BX−
Reacquisition	AX+	BX−	AX+, BX+