

UCLA

UCLA Previously Published Works

Title

Design of a Neurally Plausible Model of Fear Learning

Permalink

<https://escholarship.org/uc/item/5x45p67x>

Journal

Frontiers in Behavioral Neuroscience, 5(JULY)

ISSN

1662-5153

Authors

Krasne, Franklin B
Fanselow, Michael S
Zelikowsky, Moriel

Publication Date

2011

DOI

10.3389/fnbeh.2011.00041

Peer reviewed



Design of a neurally plausible model of fear learning

Franklin B. Krasne^{1,2*}, Michael S. Fanselow^{1,2,3} and Moriel Zelikowsky¹

¹ Department of Psychology, University of California Los Angeles, Los Angeles, CA, USA

² Brain Research Institute, University of California Los Angeles, Los Angeles, CA, USA

³ Department of Psychiatry and Biobehavioral Sciences, University of California Los Angeles, Los Angeles, CA, USA

Edited by:

Jeansok J. Kim, University of Washington, USA

Reviewed by:

Stephan G. Anagnostaras, University of California, San Diego, USA

Gavan McNally, University of New South Wales, Australia

*Correspondence:

Franklin B. Krasne, Department of Psychology, University of California Los Angeles, Los Angeles, CA 90095, USA.
e-mail: krasne@psych.ucla.edu

A neurally oriented conceptual and computational model of fear conditioning manifested by freezing behavior (FRAT), which accounts for many aspects of delay and context conditioning, has been constructed. Conditioning and extinction are the result of neuromodulation-controlled LTP at synapses of thalamic, cortical, and hippocampal afferents on principal cells and inhibitory interneurons of lateral and basal amygdala. The phenomena accounted for by the model (and simulated by the computational version) include conditioning, secondary reinforcement, blocking, the immediate shock deficit, extinction, renewal, and a range of empirically valid effects of pre- and post-training ablation or inactivation of hippocampus or amygdala nuclei.

Keywords: fear, learning, conditioning, extinction, model, neural circuit, amygdala, hippocampus

INTRODUCTION

Cues that predict the imminent onset of pain or danger as well as the situational contexts in which such events occur come to evoke fear. In laboratory animals this has a variety of manifestations including freezing, cardiac responses, and analgesia (Bolles and Fanselow, 1980). Such fear conditioning provides one of neuroscience's most promising and active arenas for analyzing neural mechanisms of learning, generally. There is now considerable evidence that the acquisition of conditional fear can be substantially attributed to forms of long-term potentiation at synapses of cortical, thalamic, and hippocampal afferents on cells of the basolateral region of the amygdala (BLA; Blair et al., 2001). Synaptic plasticity within the central nucleus, which contains the amygdala's fear-producing output neurons and which receives input derived from BLA, also seems to contribute (Pare et al., 2004; Wilensky et al., 2006; Ciochi et al., 2010). And although there is not yet consensus on the locus of extinction, a variety of lines of work, both behavioral and neurophysiological, have established that under most circumstances extinction is at least in part due to learned inhibition of activity in fear-producing pathways and not primarily to erasure of the changes that caused the original conditioning (Bouton, 1993, 2004). While these are major advances, there are many conspicuous properties of fear learning for which we neither know the mechanisms nor have hypotheses of a kind that invite physiological testing.

The well-known phenomenon of "renewal" provides an excellent example (Bouton, 1993, 2004). It is generally agreed that once fear is conditioned to a cue, fear will be expressed when that cue is presented in most any context. However, if fear is then extinguished, either in the context of original training or elsewhere, the extinction will be quite specific to the context in which it occurred, and the CS will still evoke fear in other contexts; i.e., fear will be "renewed." Renewal provides some of the best evidence of the inhibitory nature of extinction, but hypotheses as to its mechanisms are not usually neural in character. Discussions of mechanism are generally in terms of "occasion setting," "disambiguation," "storage and

retrieval," "gating" of inhibition, etc. These are essentially psychological constructs and do not provide much guidance for implementation at a neural level.

The goal of the present work was therefore to construct a conceptual and computational *neural* model of fear conditioning that can account for what we consider some of the most salient aspects of fear conditioning and that can be justified by known neurophysiology. This model is referred to as "FRAT" (for "Fraidy Rat" after an earlier, simpler version of the model used for teaching). While for modeling purposes we focused on freezing, the model is expected to generalize to other commonly used measures such as fear-potentiated startle, autonomic arousal, conditioned suppression, and inhibitory avoidance, with the greatest differences being in response generation circuitry downstream from the BLA (Hitchcock and Davis, 1986; LeDoux et al., 1988; Tomaz et al., 1993; Amorapanth et al., 1999). The phenomena that the model was explicitly designed to account for are listed in **Table 1**. Models of this kind, which are constructed to provide possible explanations of certain phenomena, are sometimes referred to as "top-down" models. This is to be contrasted with so-called "bottom-up" models in which low-level properties of a system are comprehensively incorporated into a model and the emergent behavior of the running system then examined. A bottom-up model of fear conditioning that shares features in common with the present top-down model has recently been described and analyzed (Li et al., 2009). It incorporated much-more detailed information on synaptic and cellular properties than does the present model. It simulated both conditioning and extinction, but its scope did not permit examination of many of the phenomena of **Table 1**. The two types of models have complementary utility. Bottom-up models allow one to explore the completeness of one's understanding of the properties of a system that has been extensively analyzed; if one really understood it all and incorporated it into the model, the model ought to successfully emulate the functioning system. Top-down models allow one to explore what sorts of lower-level properties and inter-relations could explain

Table 1 | Design targets.**CONDITIONING**

- A. Conditioning contingent on CS onset preceding US onset.
- B. Conditional stimuli should reinforce new fear learning (higher-order conditioning).
- C. Cue conditioning and context conditioning should be stimulus-specific.
- D. CS fear should be well expressed outside the training context¹.
- E. In an unfamiliar context the US must be delayed if context fear conditioning is to occur (“immediate shock deficit”)².
- F. Blocking should occur and be PAG opiate receptor-dependent³.

EXTINCTION

- G. Extinction of CS fear and context fear should be stimulus-specific.
- H. Extinction of cued fear should not generalize to new contexts (renewal) and renewal should have “gating-like” properties⁴.
- I. Extinction should be PAG opiate receptor-dependent⁵.

AMYGDALA

- J. Conditioning should require BLA synaptic plasticity^{6,10}.
- K. Cue conditioning should ordinarily be LA-dependent⁷ (but see **Table 4**).
- L. CEm should be required for expression of both cue and context fear¹⁶.
- M. Post-training ablation of BL should abolish previously learnt cued fear, but pre-training BL ablation should not prevent cue fear conditioning⁸.
- N. Context conditioning should ordinarily be BL-dependent⁹.
- O. Extinction should require BLA synaptic plasticity¹⁰.

HIPPOCAMPUS

- P. Newly acquired contextual fear should be lost if hippocampus-ablated (retrograde amnesia) but long established (“remote”) contextual fear should not¹¹.
- Q. Pre-training hippocampal suppression (not ablation) or NMDAR blockage should preclude context fear learning (anterograde amnesia)¹².
- R. Cued fear learning should ordinarily be hippocampus-independent¹¹.
- S. Context conditioning to long-familiar contexts is not hippocampus-dependent (no retrograde or anterograde amnesia)¹³.
- T. After pre-training hippocampal ablation, extinction should be context-specific as evaluated by ABA, ABC, and AAB renewal¹⁸.
- U. Pre-training hippocampal ablation should not prevent context fear learning but should accentuate the immediate shock deficit¹⁴.
- V. After hippocampal ablation, acquisition of context conditioning should be PFC-dependent¹⁵.
- W. Expression of remote context fear should be PFC-dependent¹⁷.

*References*¹Bouton and King (1983).²Fanselow (1990).³Cole and McNally (2007).⁴Bouton (1993, 2004), Harris et al. (2000).⁵McNally et al. (2004b), Parsons et al. (2010).⁶Rodrigues et al. (2001).⁷Goossens and Maren (2001), Nader et al. (2001), Calandreau et al. (2005).⁸Anglada-Figueroa and Quirk (2005), Jimenez and Maren (2009).⁹Helmstetter (1992), Fanselow et al. (1994), Yaniv et al. (2004), Calandreau et al. (2005), Onishi and Xavier (2010).¹⁰Sotres-Bayon et al. (2007), Zimmerman and Maren (2010).¹¹Kim and Fanselow (1992).¹²Young et al. (1994), Stiedl et al. (2000), Bast et al. (2003), Quinn et al. (2005), Parsons and Otto (2008), Schenberg and Oliveira (2008), Raineke et al. (2010).¹³Young et al. (1994), Anagnostaras et al. (2001).¹⁴Wiltgen et al. (2006).¹⁵Zelikowsky et al. (2010).¹⁶See Wilensky et al. (2006), Ciochi et al. (2010).¹⁷Quinn et al. (2008a).¹⁸Wilson et al. (1995), Frohardt et al. (2000), Zelikowsky et al. (2011) but see Ji and Maren (2005).

the known behavior of a complex system; they can guide one in the search for actual mechanisms and the interpretation of existing observations.

The present model can account for many properties of Pavlovian fear conditioning. Most of the ideas upon which it was based have been separately expressed by various authors. However, the inclusion of a computational implementation makes it possible to determine whether and how independently plausible mechanisms can work together in a comprehensive model. It is to be expected that the attempt to engineer a system that can successfully emulate the behavior produced by biological fear-learning circuitry will help us to make sensible interpretations of experimental findings on actual fear-learning circuitry and will aid efforts to determine the actual neural mechanisms of fear conditioning and extinction.

MATERIALS AND METHODS**SCOPE**

FRAT is an attempt to construct a relatively simple neurally oriented model of context conditioning and of “delay” cue conditioning, in which CS and US offset coincide or the US begins contiguously with CS offset. FRAT is not presently subject to trace conditioning.

The model has been made computational because we believe informal reasoning does not provide a reliable way of determining how the fairly large number of mechanisms needed to account for observed phenomenology will interact. Accurate qualitative predictions are the model’s primary goal. Although some ordinal quantitative predictions of the model may be useful, available information is insufficient to guide production of a model that could be expected to make comprehensive quantitative predictions.

The phenomena that FRAT was constructed to emulate are those listed in **Table 1**. There are certainly some phenomena, such as latent inhibition and trace conditioning, that the model does not presently simulate, nor intend to. Although we believe that small additions would have allowed simulation of these and other phenomena, we have wished to limit the model’s complexity. However even as it stands, the model makes a substantial number of empirically valid predictions that go beyond the design-criteria of **Table 1** and many predictions that remain to be tested.

At present FRAT deals with only two CSs (CS1 and CS2) and three contexts (A, B, and C), which are considered to have completely non-overlapping neural representations. This has allowed

us to defer issues of storage capacity and related complexities that are better initially considered independently of the matters listed in **Table 1** and that we hope to address in later work.

FRAT NEURONS

The resting potential of FRAT neurons is, for convenience, taken as zero. Membrane potential (V) is specified as deviation from rest. Active synapses open postsynaptic ion channels (cause conductances) for particular ions. Excitatory input causes ion movements that depolarize the membrane potential toward an excitatory reversal potential E (taken as 100 mV above rest). Inhibitory input opens ion channels that have an equilibrium potential near the resting level and thus result in current flows that move the membrane toward the resting level (as with GABA-mediated chloride conductance in real neurons in cases where the neurons are not persistently depolarized by tonic excitatory input). Therefore, inhibitory input, rather than working by hyperpolarizing the cell, is effective mainly because it allows excitatory currents to pass out of the cell instead of depolarizing it, thus attenuating EPSPs (so-called “divisive” inhibition).

When a FRAT neuron becomes depolarized beyond its firing threshold, its firing rate (also referred to as “activity” or “activation”) increases. Maximum activity, taken as unity, is reached at a depolarization level that varies according to neuron type. Individual spikes are not represented in FRAT, only firing rates.

SYNAPTIC STRENGTH AND SYNAPTIC CONDUCTANCE

Synapses can vary in strength or synaptic “weight” (W) according to their type and history of stimulation. The greater a synapse’s weight, the more postsynaptic conductance G (conductances are expressed relative to leakage conductance) will be produced by a given level of presynaptic activity A according to the relationship,

$$G = A \cdot W \quad (1)$$

SUMMATION OF EXCITATION AND THE INTERACTION OF EXCITATION AND INHIBITION

The neurons of this model treat the summation of excitatory inputs as though all excitatory input synapses were located close together on their postsynaptic targets; in computational terms, they all innervate a single “dendritic compartment.” Inhibition has quite different consequences depending on whether inhibitory synapses are located electrically close to excitatory synapses (i.e., in the same electrical compartment) or at a separate location more proximal to the soma (for modeling purposes, a second electrical compartment); we thus distinguish here between “distal” and “proximal” inhibition (Vu and Krasne, 1992). When inhibition is distal, it competes more or less symmetrically with excitation; an increase in one can always be offset by an increase in the other. However this is not the case for proximal inhibition, which for reasons explained below, was used here.

Denoting the sum of all the conductances generated by excitatory input as G_e , the equilibrium potential for excitation as E , and the sum of inhibitory conductances by G_i , the depolarization produced by summed inputs is of the form

$$V = \left(\frac{G_e \cdot E}{1 + G_e} \right) \left(\frac{1}{1 + G_i} \right) \quad (2)$$

In the absence of inhibition, membrane potential is given by the ratio in the left brackets, which is graphed in **Figure 1** (bold solid line). When G_e is much smaller than unity, the depolarizations produced by each active, effective synapse sum approximately linearly. However, as G_e grows, each increment of input adds less to V , and the total depolarization becomes asymptotic to E . This non-linearity of summation is used to advantage in the model. As seen in the right bracket term of Eq. 2, proximal inhibition acts as a divisor for excitation. Its effect is shown in the thin solid lines of **Figure 1** (for comparison the effect of distal inhibition is shown by the dashed lines).

COMPUTATIONAL STEPS AND TIME RESOLUTION

A detailed technical account of FRAT is provided in Section “Supplementary On-line Materials.” However a few basic features that will be needed for what follows must be mentioned here: The state of FRATs variables is updated every 1 s of real time, and all graph abscissas are in seconds. Values of variables should be thought of as averages or peak values for the interval in which they are determined. Thus phenomena that occur on millisecond time scales are not emulated as such. Events occurring in a given, 1 s time interval cause plastic changes that are expressed behaviorally in the next.

PARAMETERS

Finding a single set of parameters (thresholds, change rates, etc.) that will allow FRAT to behave so as to meet the specifications of **Table 1** required considerable experimentation. Search routines such as Simulated Annealing might have been useful, but we chose informal experimentation because it was helpful in providing insight into the interactions of various processes. The parameters that must be set include potentiation and depression rates for BLA principal cells and inhibitors in the presence of reinforcement and extinction modulator, the relative numbers of neurons carrying each elementary and configural representation of hippocampus and cortex, the

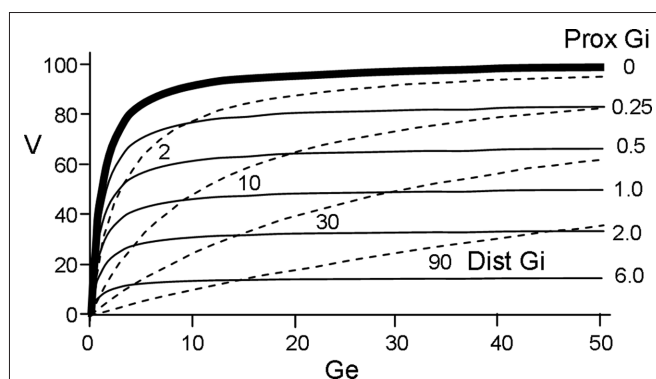


FIGURE 1 | Relationship between depolarization (V) and excitatory and inhibitory conductances (G_e and G_i). Solid curves plot depolarization as a function of G_e for various values of G_i , assuming proximal inhibition (V for distal inhibition is given by dashed curves).

thresholds and saturation depolarizations of all the neuron types indicated in **Figure 5A**, the rate of formation of hippocampal and cortical configural representations, and depolarization and “calcium” levels at which potentiation and depression began and ended in the curves of **Figure 5C**. The parameters eventually settled on (specified in section Supplementary On-line Materials) are certainly not claimed to be in any way a “best” set; however, they do establish that a model with the structure employed can generate the desired set of phenomena, which is the purpose of this paper.

NOTATION

Neuron type names are in bold font; thus, “**CEm**” refers to neuron type CEm, which may be thought of as a single neuron or as a population of neurons all of which behave similarly. “CEm,” not in bold font, refers to the central nucleus CEm.

SUPPLEMENTARY ON-LINE MATERIALS

A full mathematical description of the model and presentation of simulations for which there is not space here can be found at <http://escholarship.org/uc/item/3hk369c4>. A document providing Matlab code for FRAT and instructions for using the program is in preparation; when completed it will be placed at this URL.

RESULTS

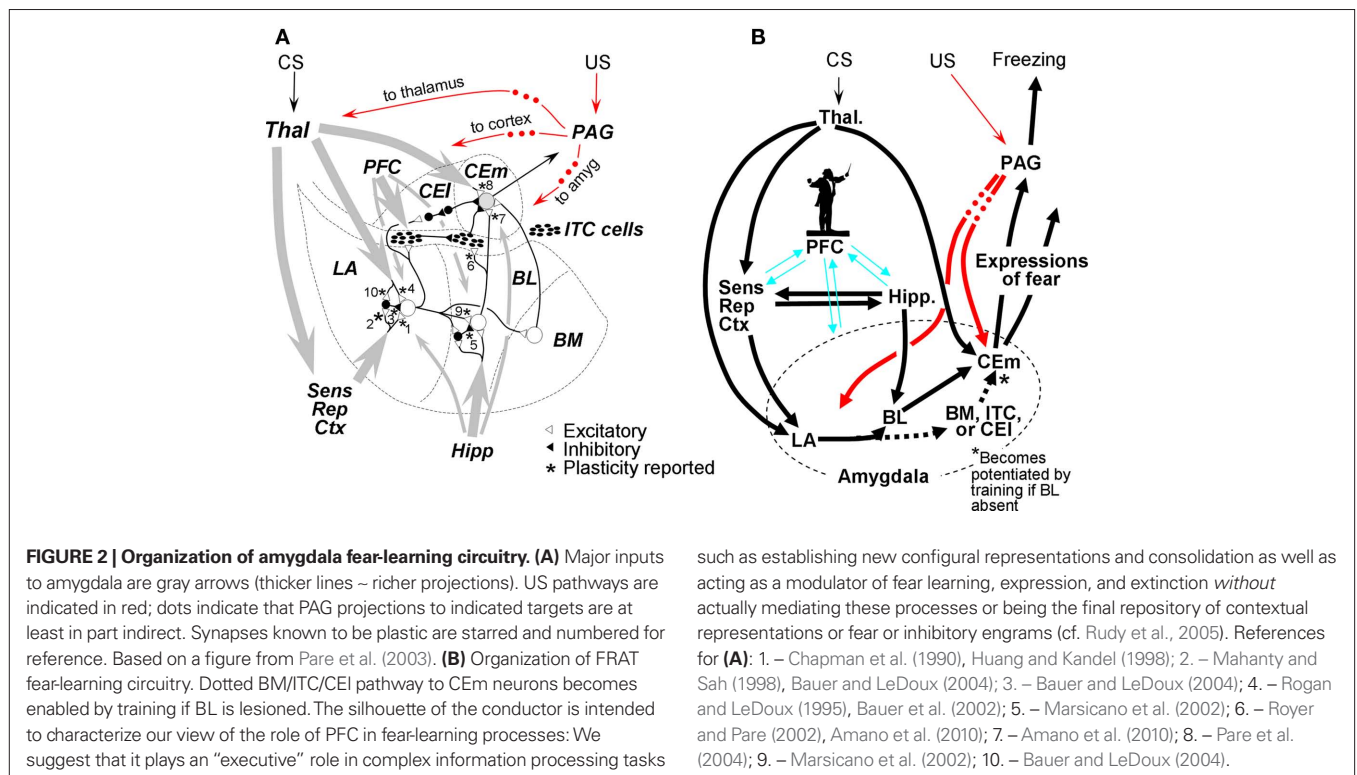
THE MODEL AND RATIONALE

In order to create a model, specific assumptions about fear learning and its circuitry had to be adopted. Many of the assumptions made are widely accepted as valid. However, some remain matters of debate. In these cases we selected particular possibilities as working hypotheses; as new facts come to light, revisions will surely be required.

Outline of major information flow pathways

Here we consider the broad BLA as composed of lateral amygdala (LA), the lateral basal area (BL), and the medial basal region (BM). These regions project via a variety of routes to medial central nucleus (CEm), which is believed to be the portal to fear-expression circuitry that lies downstream of the amygdala itself. Many of the synapses within the BLA and its pathways to CEm are plastic (starred in **Figure 2A**), and most have at one time or another been suggested to play some role in fear conditioning or extinction. Although it is generally agreed that changes mediating cue and context fear conditioning occur within BLA, the roles of the different sub-regions of the BLA are still being resolved, as is the route of information flow from LA and BL to CEm. **Figure 2A**, modeled after a figure of Pare et al. (2003), summarizes aspects of the known circuitry; **Figure 2B** portrays the circuit as implemented in FRAT.

Pathway for CS (cued) fear. Based on available data (see **Figure 2A**; McDonald, 1998; Goossens and Maren, 2001; Nader et al., 2001; Calandreau et al., 2005), the LA of FRAT receives input from cortical sensory and sensory association areas (hereafter referred to collectively as “stimulus representation” or “representation” cortex) and from thalamus (McDonald, 1998; Doron and Ledoux, 1999), and LA mediates a major portion of cued fear conditioning. Current belief is that LA output might reach CEm via BL or BM (Pare et al., 1995, 2003; Pitkanen et al., 1995; Savander et al., 1995; Royer et al., 1999) or by disinhibitory mechanisms via intercalated cells (ITC) or the lateral part of the central nucleus (CEI; Pare et al., 2004; Cioocchi et al., 2010; Haubensak et al., 2010), but since post-training lesions of BL appear to abolish conditional responses to discrete cues (Anglada-Figueroa and Quirk, 2005), FRAT LA



such as establishing new configural representations and consolidation as well as acting as a modulator of fear learning, expression, and extinction *without* actually mediating these processes or being the final repository of contextual representations or fear or inhibitory engrams (cf. Rudy et al., 2005). References for **(A)**: 1. – Chapman et al. (1990), Huang and Kandel (1998); 2. – Mahanty and Sah (1998), Bauer and LeDoux (2004); 3. – Bauer and LeDoux (2004); 4. – Rogan and LeDoux (1995), Bauer et al. (2002); 5. – Marsicano et al. (2002); 6. – Royer and Pare (2002), Amano et al. (2010); 7. – Amano et al. (2010); 8. – Pare et al. (2004); 9. – Marsicano et al. (2002); 10. – Bauer and LeDoux (2004).

was made to ordinarily access CEM via BL. However, to reflect the fact that CS fear can be acquired and extinguished normally after pre-training lesions of BL (Nader et al., 2001; Yaniv et al., 2004; Anglada-Figueroa and Quirk, 2005; Calandreau et al., 2005; Onishi and Xavier, 2010; but see Goossens and Maren, 2001), conditioning of FRAT in the absence of BL was made to cause development of an alternate pathway from LA-to-CEM.

Synaptic changes within LA are widely regarded as essential for development of cued fear. There is limited evidence that has been taken to suggest that most LA principal cells receive cortical and thalamic innervation that would allow conditioning to almost any CS (Repa et al., 2001; Rumpel et al., 2005; Han et al., 2007; Won and Silva, 2008; Zhou et al., 2009). Although far from established, we take this as justification for the initial, simplifying assumption that each LA principal cell is innervated by all afferents to LA and that after conditioning all LA principal cells respond similarly (referred to here as “uniform innervation”). Under this assumption, any synaptic changes downstream of LA would alter responses to all stimuli indiscriminately, which would lead to arbitrary and unwanted stimulus generalization. Therefore conditioning to a CS in FRAT has been made entirely due to plasticity at the level of BLA. However, in reality central nucleus plasticity appears to contribute to conditioning (Wilensky et al., 2006; Ciochi et al., 2010). In fact, if one lifts the assumptions of uniform innervation at the level of LA, plasticity downstream of BLA becomes a possible asset, as addressed in the Section “Discussion.”

Pathway for conditioning to context and other configural stimuli. Since pre-training inactivation of BL as well as suppression of activity or plasticity of hippocampus, which projects richly to BL, substantially reduces context fear conditioning (Fanselow and Kim, 1994; Young et al., 1994; Stiedl et al., 2000; Bast et al., 2003; Yaniv et al., 2004; Calandreau et al., 2005; Quinn et al., 2005; Parsons and Otto, 2008; Schenberg and Oliveira, 2008; Onishi and Xavier, 2010; Raineki et al., 2010; but see Goossens and Maren, 2001; Matus-Amat et al., 2004), the hippocampus–BL pathway was taken to mediate fear conditioning in normal (unlesioned) FRAT, and BL was assumed to directly excite CEM. The actual pathway to BL in rodents may be dorsal hippocampus–ventral hippocampus–BL (see Fanselow and Dong, 2010), but dorsal and ventral hippocampus were not distinguished in the model.

The role commonly attributed to the hippocampus in context fear learning is the rapid exposure-dependent establishment of neural representations of elemental stimulus conjunctions, including contexts, independent of actual training with a US (Kiernan and Cranney, 1992; Fanselow, 2000; Rudy and O’Reilly, 2001; Rudy, 2009). It is presumed to be because of this that at least some tens of seconds of exposure to a context must precede US onset for context fear to be learned (the so-called “immediate shock deficit”; Blanchard et al., 1976; Fanselow, 1986, 1990; Landeira-Fernandez et al., 2006). This essential rapid learning process is prevented if the activity or plasticity of the dorsal hippocampus is suppressed (Matus-Amat et al., 2004; Stote and Fanselow, 2004). These characteristics were design-criteria for FRAT.

It is often postulated that hippocampus re-establishes in cortex representations of the full set of elemental cues that compose a context (e.g., Marr, 1971; McClelland et al., 1995). FRAT contains

circuitry designed to implement such reinstatement. However, if reinstated representations of elementary contextual cues provide sufficient input to amygdala to support context fear conditioning, then according to the anatomy proposed in **Figure 2B**, context fear learning would remain possible despite BL lesions. Since we decided above to accept those data which suggest that such learning does not occur in rats, projection to amygdala of reinstated elementary cues has been made of negligible strength in the simulations of this paper. Nevertheless it must be acknowledged that BL lesion studies do typically find some residual context conditioning; therefore the consequences of increasing the effect of reinstated contextual cues should at some point be explored.

Alternative representations of configural stimuli- consolidation and compensation. It is well-known that whereas context fear is lost if hippocampus is ablated soon after conditioning (retrograde amnesia), this does not happen if sufficient time passes before ablation (Kim and Fanselow, 1992). It is believed that conjunctive representations established in the hippocampus are gradually shifted to cortex, the inferred process of “consolidation” (our references here will be only to this slow, “system” consolidation as opposed to the much faster, protein synthesis-dependent stabilization of synaptic change sometimes referred to as “cellular” consolidation). Moreover, if sufficient time for consolidation is allowed following simple familiarization, before context conditioning is attempted, no retrograde amnesia results from post-training hippocampal ablation (Anagnostaras et al., 2001). Thus, hippocampal representations, once formed by incidental learning, seem to become re-established in cortex even in the absence of conditioning. These characteristics were design requirements for FRAT (**Table 1**). Consolidation was assumed to cause replacement of hippocampal representations of familiarized conjunctive stimuli with representations located in sensory-representation cortex. Since hippocampus innervates BL, whereas representation cortex innervates LA, post-conditioning consolidation must not only transfer conjunctive representations from hippocampus to cortex but must also replace hippocampus–BL synaptic changes with cortex–LA changes that will produce similar behavior.

Whereas suppression of hippocampal function greatly impairs context conditioning, actual removal of hippocampus does not (Maren et al., 1997; Frankland et al., 1998; Cho et al., 1999; Richmond et al., 1999; Rudy et al., 2002; Wiltgen et al., 2006). One interpretation of this has been that intact hippocampus suppresses conditioning of fear to cortical representations of the elemental cues that compose contexts (Matus-Amat et al., 2004; Biedenkapp and Rudy, 2009; Rudy, 2009), and thus, once hippocampus is removed, conditioning to elemental cues of context becomes possible, whereas before it was not. However, if cortically represented elements of context that were present but suppressed before hippocampal removal, and operative after hippocampal removal, were capable of supporting context conditioning, then they should be able to support context conditioning even without any pre-exposure to context. However, the fact appears to be that even hippocampus-ablated rats require tens of seconds of experience with a context before they can be conditioned to it, just as do normal rats (Wiltgen et al., 2006). This suggests that conditioning in such animals is not to pre-existing elemental cues, but rather

it is to conjunctive representations that, in the absence of hippocampus, form cortically over tens of seconds as they do within the hippocampus in normal animals (Fanselow, 2010). We refer to this hypothetical process as cortical “compensation.” In FRAT such compensation occurs and creates conjunctive representations in sensory-representation cortex; under the parameter values used for the simulations of this paper, elemental cues of context are never the basis for context conditioning.

It is not surprising that more signs of cortical compensation should be seen after hippocampal ablation than during disruptions of function, since ablation causes degeneration as well as dysfunction. However, whether some degree of compensation also occurs when only function is disrupted, is unclear. Limited context fear is often, though not always learned despite hippocampal APV or muscimol infusion (Young et al., 1994; Stiedl et al., 2000; Bast et al., 2003; Quinn et al., 2005; Parsons and Otto, 2008; Schenberg and Oliveira, 2008; Raineki et al., 2010), but it is not known whether such partial effects should be attributed to compensation or to manipulations that were incomplete. As a working hypothesis for the design of FRAT, we assumed the latter; in FRAT no compensation is produced by suppression of hippocampal activity or plasticity unless hippocampus is actually removed. This choice led to an unexpected ability to correctly predict the outcome of several experiments on hippocampal involvement in extinction described below.

Assumptions about the development and locus of conjunctive representations were important in the construction of FRAT not only because of the relevance of such representations to learning of context fear, but also because inhibition associated with such cues can contribute to the context-specificity of extinction that is revealed by renewal. Therefore, if as just discussed, hippocampus-ablated animals can form conjunctive representations cortically, it might be expected that they would also be subject to renewal. The literature on renewal in hippocampus-ablated rats is somewhat conflicting (Wilson et al., 1995; Frohardt et al., 2000; Ji and Maren, 2005; Zelikowsky et al., 2011); however we felt that there was sufficient evidence that such rats are subject to fairly normal renewal to make this capacity one of the design-criteria of **Table 1**.

Pre-frontal cortex. The pre-frontal cortex (PFC), which is reciprocally well connected to amygdala and hippocampus (**Figure 2**), affects fear conditioning in several ways:

1. Although, as discussed above, context conditioning occurs almost normally when hippocampus is ablated, this is not the case if PFC (infralimbic or prelimbic) has also been ablated (Zelikowsky et al., 2010).
2. Even with hippocampus intact, context-conditioned IL-PFC-damaged animals that have been trained to fear one context may also treat a neutral context as dangerous (Zelikowsky et al., 2010). Such over-generalization could be taken to suggest that PFC is involved in normal hippocampal pattern separation, and thus perhaps in the normal operation of the hippocampus, generally.
3. Consolidated context fear is reduced by post-consolidation PFC lesions (see **Table 1**, W; Frankland and Bontempi, 2005; Quinn et al., 2008a; Nieuwenhuis and Takashima, 2011). Given

our assumption that a representation cortex-to-LA pathway replaces the hippocampus-to-BL pathway of original learning, this suggests that full activation of cortical contextual representations may require PFC.

4. In animals with certain PFC lesions, extinction is not well retained unless extra extinction training is given, though extinction itself seems to occur normally (Quirk et al., 2000; Lebron et al., 2004). It has also been found that pre-frontal neurons can recruit inhibitory activity within amygdala (Milad and Quirk, 2002), and that various drugs thought to affect learning or consolidation prevent or retard the stabilization of extinction when applied to PFC (Santini et al., 2001, 2004; Burgos-Robles et al., 2007). These and related observations are sometimes taken to suggest that extinction occurs because learning within PFC causes it to recruit inhibition of amygdala circuitry. However, extinction itself is normal in PFC-lesioned animals; it is only its persistence that is altered. Moreover, once extinction has occurred, inactivation of those parts of the PFC thought to drive inhibition of amygdala fear circuitry do not cause a return of fear responses (Sierra-Mercado et al., 2006); and drugs that are thought to affect learning or consolidation prevent extinction when applied to amygdala itself as well as to PFC (e.g., Lin et al., 2003; Herry et al., 2006; Zimmerman and Maren, 2010). Thus, while there is no doubt that PFC affects the stabilization of extinction, the exact nature of the effect is not yet understood.

The PFC is commonly postulated to play some sort of executive role that enables or facilitates complex information processing tasks that extend over time (Miller and Cohen, 2001). We think the potent, but somewhat complex and ill-defined effects of PFC on fear are consistent with its acting as a modulator of fear learning, expression, and extinction without it actually mediating these processes or being the final repository of contextual representations or fear engrams. Moreover, given the significance of context in extinction learning, we believe that the role of PFC in extinction may well be, at least in part, secondary to the role in context learning suggested by points 1–3 above (see Morgan et al., 2003; Zelikowsky et al., 2010). Indeed, if the normal functions of the hippocampus are in fact somewhat PFC-dependent, then one possible interpretation of the loss of extinction during the first 24 h after its establishment in PFC-lesioned rats is that hippocampal pattern-completion mechanisms are compromised. These rats may fail to compensate for small changes in context due to the passage of time, with the result that extinction, which is extremely context-specific, suffers more renewal due to the passage of time than would normally be the case. What appears as “forgetting” might merely be such renewal.

Given the above observations, and considering the presumed executive function of PFC, both the processes of compensation and consolidation and the activation of cortical conjunctive representations in FRAT have been made PFC-dependent. Some precedent for a role of PFC in the consolidation process is provided by recent evidence that off-line PFC activity is required for post-training stabilization of motor skill learning (Kantak et al., 2010).

Pathway for consolidated and compensated contextual fear. The destination within amygdala of cortical neurons that, through compensation or consolidation, have come to represent cue conjunctions

have not been studied. Sensory association areas preferentially innervate LA, but very high level association cortex, and especially polymodal cortex, may also innervate BL (McDonald, 1998; Pitkanen et al., 2000; Lindquist et al., 2004; Suzuki, 2010). In FRAT we made these cortical representation neurons project exclusively to LA.

One can only guess whether the same or different cortical neurons contribute to the representations formed as the result of compensation and consolidation. It is commonly supposed that consolidation causes the development of cortical representations that are extensively integrated with other pre-existing cortical representations. Representations formed by compensation are perhaps more akin to temporary working memories or hippocampal “snapshots” of cortical activity patterns (as in Marr, 1971). Representations formed by consolidation and compensation might therefore be expected to be separate and have different properties. However, pending evidence of this, we provisionally opted for simplicity and used the same cortical neurons for both types of representation in FRAT.

Table 2 lists the main assumptions discussed above.

US Pathway. There are a great many anatomical routes via which information about footshock USs might reach the baso-lateral and central complexes of the amygdala. Nociceptive information reaches the brain at via spino-thalamic, spino-reticular, spino-mesencephalic, spino-parabrachial, and perhaps still other pathways, and each of these can then project nociceptive information to the amygdala via both direct and indirect routes (see, e.g., Willis and Westlund, 2004). Given this abundance, it has been difficult to define which pathways are actually necessary or sufficient to reinforce fear learning (Shi and Davis, 1999; Brunzell and Kim, 2001). However, a recent study (Johansen et al., 2010b) has reported that muscimol inactivation of periaqueductal gray (PAG) of the spino-mesencephalic pathway, electrical stimulation of which can act as a US for fear conditioning (Di Scala et al., 1987), abolishes at least cued fear conditioning. PAG output can itself reach amygdala via intralaminar thalamus, anterior cingulate, hypothalamus, locus coeruleus, and ventral tegmental area, and, again, which pathways are most important has not been determined (see, e.g., Herrero et al., 1991; Shi and Davis, 1999; Brunzell and Kim, 2001; Lanuza et al., 2004; Johansen et al., 2010b). For the purposes of designing FRAT, we accept the likely importance of portions of the PAG as an ascending route for the US, in addition to its being on the output pathway mediating freezing (see Carrive, 1993), but we do not try to specify further.

The fear-expression model

The operation of the model once learning has occurred is described first:

Conditioning is due to potentiation of thalamic, cortical, and hippocampal – BLA synapses. Conditioning is due to potentiation of synapses made by cortical and thalamic afferents on LA principal cells and hippocampal afferents on BL principal cells. LA input mostly codes simple cues used as CSs, but may also code for complex configurations (as in visual configuration-sensitive cells of inferotemporal cortex (e.g., Miyashita and Chang, 1988; Perrett et al., 1992) if these have developed as the result of consolidation or compensation (see previous section), or other processes. BL principal cells are innervated by LA output cells via non-plastic synapses as well as by conjunction-representing neurons of hippocampus.

Extinction is due to potentiation of afferent-LA and afferent-BL inhibitory interneuron synapses. The facts of renewal demand that extinction be at least in part due to active inhibition of responses whose ability to be excited is not erased. Given the assumption of uniform innervation (see above), extinction can only be stimulus-specific if inhibition is recruited via synapses made by LA and BL afferents. In FRAT these are synapses of LA and BL afferents on feed-forward inhibitory interneurons (Hestrin and Galarreta, 2005; Woodruff and Sah, 2007).

A recent paper by Herry et al. (2008) has reported a class of projection neurons in BL (referred to there as “extinction” neurons) that increase their responses during extinction, while activity of another class that previously became active during conditioning (“fear” neurons), declines. This suggests that in BL, afferents may at least in part recruit inhibitory interneurons via a special class of projection neurons. Computationally this would be equivalent to what was assumed here, though the downstream effects of the projecting extinction neuron activity, which are unknown, presumably have consequences of their own.

BL receives crucial input from Context/CS conjunction-representing neurons. The phenomena of renewal require that the inhibition that mediates extinction of a CS be preferentially operative in the context where extinction occurred. In order to achieve this it was assumed that hippocampus forms representations of conjunctions of context and CSs that have been experienced in the context.

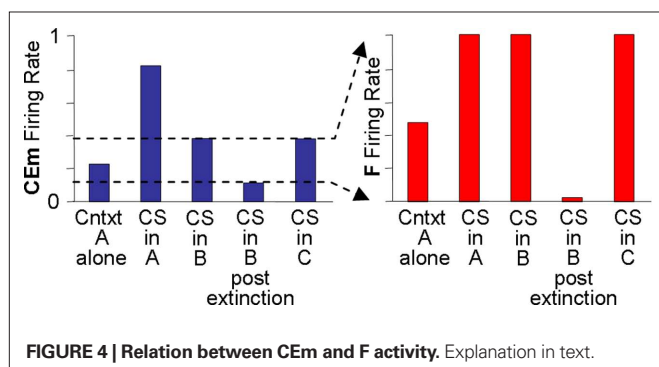
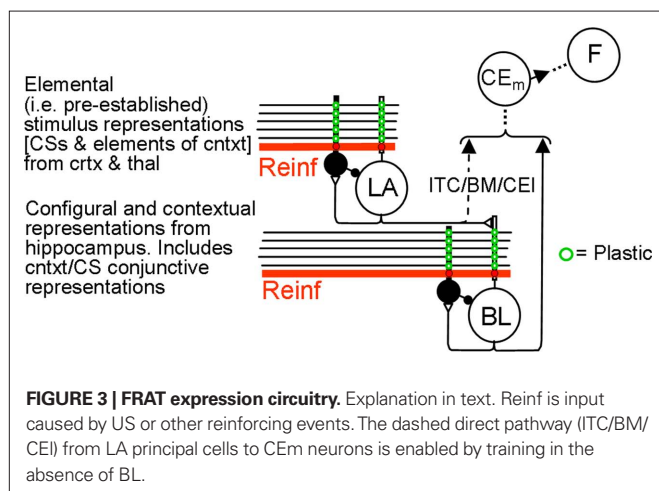
Table 2 | Summary of main assumptions.

1. Information about CSs and contexts is provided by a thalamus-sensory-representation cortex (defined in text) pathway.
2. Cued fear is normally learned in LA.
3. Conjunctive representations (including representations of context and context/cue conjunctions) are rapidly developed in hippocampus during exposure and projected to BL where fear becomes conditioned to them.
4. The normal route from LA to the output region CEm is via BL.
5. Fear extinction occurs within LA and BL.
6. Over time, hippocampal representations gradually shift to sensory-representation cortex, and fear conditioned to the hippocampal representations becomes conditioned to their cortical replacements. This process is PFC-dependent.
7. In absence of hippocampus a PFC-dependent compensation process allows sensory-representation cortex to rapidly establish conjunctive representations.
8. In absence of BL, training with a US rapidly establishes a functional pathway from LA-to-CEm via ITC, CEI, and/or BM.

Inhibition conditioned to such representations at the level of BL is the mechanism used in FRAT to produce context and CS-specific inhibition of fear responses. Blair and colleagues (Moita et al., 2003) have in fact seen hippocampal units that respond to a conditioned CS only when the rat is in a particular location.

Saliency of contextual representations is reduced by habituation. Stimuli such as contextual cues that are continuously present would be expected to be less salient and have less impact than stimuli that have sudden onset and are present only briefly (Phillips and LeDoux, 1992). This is captured in FRAT by partially discounting the activations of contextual inputs to LA and BL by an “Habituation Factor” when calculating the activity of amygdala neurons; habituation in FRAT affects only performance, not synaptic weight alterations.

Output from LA and BL is truncated by downstream fear-expression mediating neurons. Figure 4 shows the sort of output that a FRAT CEM cell (Figure 3) would produce in an experiment where training occurred in context A, extinction in B, and a test was then given in C (“ABC renewal”). After conditioning the CS in context A, potentiation would have developed at synapses of CS-representing neurons on LA principal cells, context A-representing neurons on BL principal cells, and context A/CS conjunction-representing neurons on BL principal cells. As a result, CEM neurons would respond to context A itself and to CS *in* context A as shown in the first two bars on the left of Figure 4. When moved to context B there would



be no response either to the new context by itself or to the context B/CS conjunction. However, responding to the CS itself would remain, with the result that some CEM firing in the new context would still occur, as shown in the third bar on the left. After extinction in context B, inhibition would be conditioned in varying degrees to the CS itself, to context B itself, and to the context B/CS conjunction. This would result in greatly diminished CEM response to the CS when tested in context B, as shown in the fourth bar on the left of Figure 4. After the post-extinction move to context C the only inhibition would be to the CS itself, with the result that responding to the CS would be greater than it was in B as shown in the last bar on the left. If manifestations of fear were a direct reflection of this profile of activity, there would be two deviations from reality: (1). Although the fear produced by the CS would transfer substantially from context A to context B, it would be much stronger in A than B. (2). Extinction, though asymptotic, would be incomplete. The model deals with such difficulties by interposing between CEM and the activation of fear behavior a fear-expression neuron F whose firing threshold and dynamic range are such that F fails to fire when CEM activity is below level a and fires at its maximum rate when CEM activity is above b . The firing rate of F would then be as shown in the right panel; the result of the limited dynamic range of F is that expression of conditional fear prior to extinction is the same in B as in A, and behavioral extinction in B is complete even though some activity of CEM remains.

Learning rules

Some forms of LTP have properties that seem to make them ideal as a mechanism for associative conditioning. Thus, it is natural to imagine that fear conditioning might be simply a matter of establishing LTP at CA1-like synapses of CS-representing cells on CR-producing cells by pairing presynaptic activity with strong US-caused postsynaptic depolarization. However, this simple Hebbian scenario by itself seems to us to be an insufficient basis for simulating many of the properties of real fear learning.

Determination of synaptic change at amygdala principal cell synapses. Three considerations have caused us to make potentiation of presynaptically active synapses on FRAT LA and BL principal cells dependent on something other than input which directly represents the US: (1) We do not want a CS that begins after the onset of the US to cause conditioning; but it *would* do so if LTP on principal cells were determined by direct US-caused depolarization. Although some procedures, demonstrate the possibility of simultaneous or backward pairing or even the superiority of the latter for producing associations (Mowrer and Aiken, 1954; Heth and Rescorla, 1973; Mahoney and Ayres, 1976; Rescorla, 1980; Ayres et al., 1987; Matzel et al., 1988; Albert and Ayres, 1997; Esmoris-Arranz et al., 2003; Leising et al., 2007), it appears to us that for the sorts of experiments modeled by FRAT, forward pairing produces considerably more substantial and reliable conditioned fear than do non-forward procedures when fear is measured by CS-produced suppression of an ongoing operant, aversion, or freezing (Mowrer and Aiken, 1954; Heth and Rescorla, 1973; Mahoney and Ayres, 1976; Ayres et al., 1987; Albert and Ayres, 1997; Esmoris-Arranz et al., 2003). We have chosen to approximate this situation by making the model produce freezing only after forward pairing. We have also chosen to make

this rule apply to the formation of associative change *per se*, despite the possibility that the relative ineffectiveness of non-forward conditioning could be due to a performance rather than a learning deficit (Rescorla, 1980; Matzel et al., 1988). (2) A well-conditioned CS should block conditioning of a new stimulus presented along with it (Kamin, 1969). Furthermore, blocking should occur even if the representation for one of the stimuli innervates LA while that of the other innervates BL. (3) Second-order conditioning must be possible. Moreover, a CS that causes CEm activity and fear due to potentiation that has occurred in LA (e.g., a well-conditioned cue) should be able to reinforce new conditioning mediated by potentiation in BL (e.g., to a contextual stimulus), and conversely (Marlin, 1983; Helmstetter and Fanselow, 1989).

The first requirement, was met by having the depolarization needed for LTP induction in LA and BL principal cells produced, not by input directly from the US, but rather by input from a “reinforcing” neuron R that fires only transiently and only at US onset. R fires briefly when its input *increases*, which it does when the activity of the US-representing neuron U increases (Figure 5A). The transience of US-evoked depolarizing input to principal cells is consistent with the observation in rats that BLA principal cells fire only briefly at the onset of a US (Johansen et al., 2010b).

The second requirement was met by letting CEm inhibit the U neurons, which they do via inhibitory neuron Ω . Consistent with this, responses of rat BLA principal cells to shock are markedly attenuated when an established CS precedes them (Johansen et al., 2010b) and the apparent aversiveness of shock, as measured by post-shock freezing, is diminished if the shock comes during a previously well-conditioned CS (Fanselow and Bolles, 1979b). Moreover, opiate receptor antagonists reduce both blocking and the CS-caused diminution in post-shock freezing (Fanselow and Bolles, 1979b; Cole and McNally, 2007), and the locus of these effects appears to be the PAG (Hammer and Kapp, 1986; Helmstetter and Landeira-Fernandez, 1990; Cole and McNally, 2007). For this reason, the inhibition produced by FRAT’s Ω neuron is taken to be mediated by an opioid transmitter. The quantitative effect of ongoing CEm activity on responses of R to USs is shown in Figure 5B₁. It should be noted that these mechanisms of blocking are very similar to those which appear to operate in the cerebellar circuit that mediates eye-blink conditioning (Fanselow, 1998; Kim et al., 1998).

Although introduction of the reinforcing neuron R and its inhibition by Ω may appear to allow the satisfaction of requirements (1) and (2), problems remain if R exerts its effect only by depolarizing BLA principal cells. Although, as described, onset

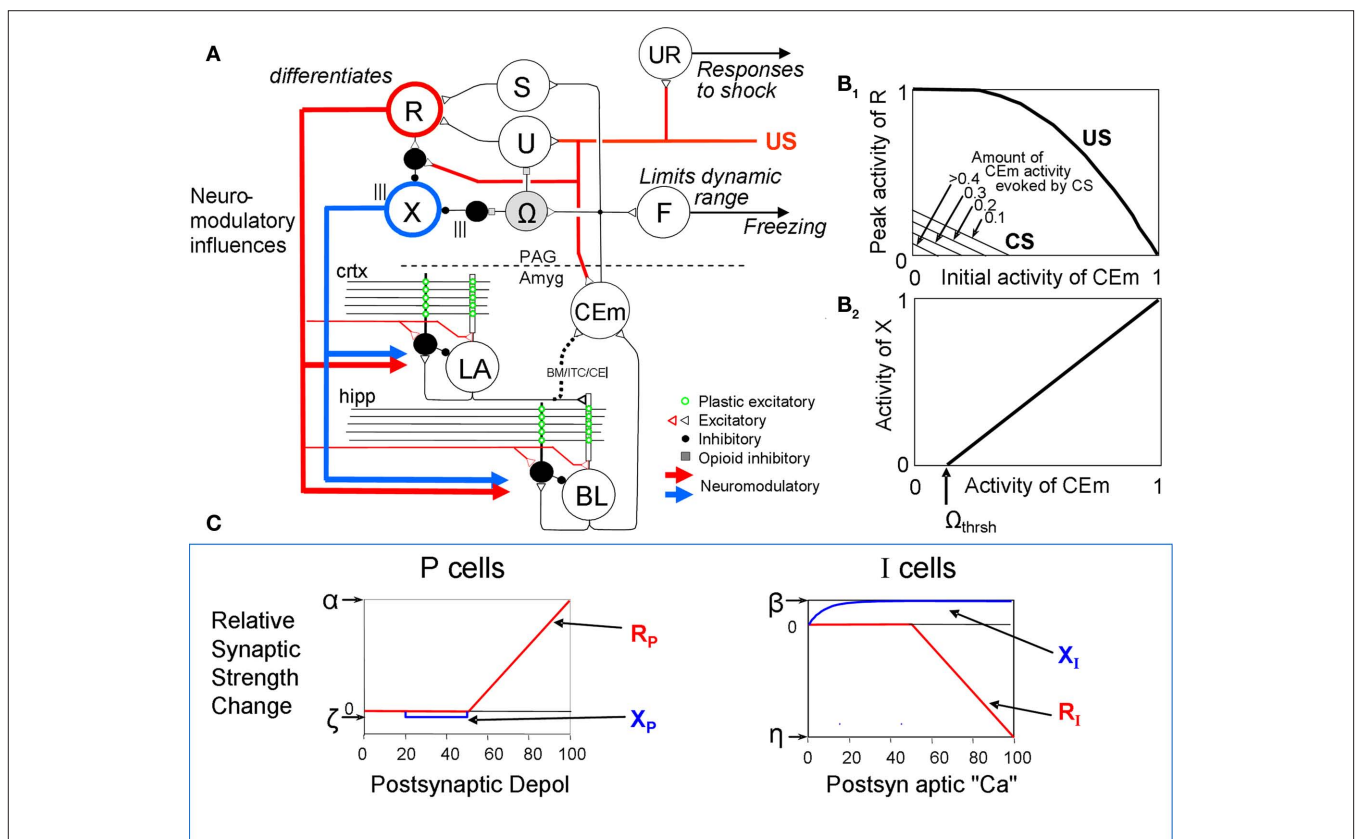


FIGURE 5 | Full FRAT circuit and plasticity properties. (A) Circuitry controlling plasticity has been added to the circuit of Figure 3. The added elements are explained in the text. **(B)** Activity of R and X as a function of CEm activity at time of reinforcer onset. The curve labeled US indicates the effect of a full-strength footshock. The curves labeled CS refer to the secondary reinforcing effect of a CS that caused the amount of CEm activity indicated. **(C)**

Neuromodulator-dependent factors determining change of synaptic strength of synapses on amygdala principal cells and inhibitory interneurons. Left: Factors R_p and X_p of Eq. 3 as a function of postsynaptic depolarization for amygdala principal (P) cells. Right: Factors X_i and R_i of Eq. 4 as a function of amount of “Ca” for inhibitory interneurons (I). Parameters α , β , η , ζ are max levels of the curves shown.

of a new CS during a continuing *unconditional* stimulus would not cause conditioning to the CS, occurrence of a new CS during continued exposure to a previously well-conditioned CS *would* cause conditioning because the well-conditioned CS would cause a maintained depolarization of principal cells due its direct input; this would violate requirement (1). Other complexities would arise during blocking experiments, especially if the US were replaced by a previously conditioned CS. Such problems were avoided in FRAT by making potentiation depend not only on principal cell depolarization but also on a neuromodulatory input provided by R. Such an assumption is far from lacking empirical precedent. Development of virtually all forms of LTP and LTD as studied in slices appear to be sensitive to most neuromodulators. The VTA dopamine system provides an obvious precedent for neuromodulatory control of learning (Schultz and Dickinson, 2000). And there is increasing evidence in rats that reinforcement of fear conditioning is substantially dependent on dopaminergic or noradrenergic modulation (Johnston et al., 1999; Bissiere et al., 2003; Bush et al., 2010; Johansen et al., 2010a).

The third requirement was met by establishing a connection between **CEm** neurons and R-cells via a “secondary reinforcement” neuron S (**Figure 5A**). This connection allows sufficiently strong conditional responses to drive R at their onset and thereby both depolarize and provide reinforcing neuromodulator to LA and BL principal cells.

Blocking of second-order conditioning (Leyland and Mackintosh, 1978) could have been achieved by letting Ω inhibit S as well as U. However, if **CEm** is already very active due to the presence of a previously conditioned stimulus when a secondary reinforcement occurs, S will be able to increase its activity very little and thus will cause very little activity of R. Consequently, blocking of higher-order conditioning should occur even without input to S from Ω , and such a connection was not included in the model. The quantitative effect of prior **CEm** activity on responses of R to secondary reinforcers is shown in **Figure 5B₁**.

Determination of change at amygdala inhibitory interneuron synapses. Extinction in FRAT is due to LTP at synapses made on feed-forward inhibitory interneurons by the same afferents that feed LA and BL principal cells. Thus, LTP of synapses on inhibitors must develop when non-reinforced CSs occur. Inhibition should be sufficient to reduce **CEm** activity below the threshold a of the freezing neurons F, but it need not totally prevent **CEm** activation. Indeed, if inhibition at the level of LA is too great, renewal will not occur when shifts of context reduce inhibition at the level of BL. On the other hand, in order to simulate the observation that extinction remains approximately normal in BL-lesioned rats, sufficient inhibition must be able to develop at the level of LA to allow extinction. By making development of LTP at synapses on inhibitors contingent on a sufficient amount of activity of both **CEm** cells and the principal cell targets of the inhibitors, it can be assured that extinction will occur under appropriate conditions without inhibition at the level of LA growing too great to allow renewal.

Control of inhibitor LTP by **CEm** activity is mediated by “extinction” neurons X (**Figure 5A**), which distribute a neuromodulator that enables LTP development on inhibitory neurons and has other extinction-promoting effects. Although it would have been simplest

to let **CEm** neurons directly excite X-cells, it is the case that in rats extinction is prevented by PAG opiate antagonists (McNally et al., 2004b; Parsons et al., 2010); therefore in FRAT **CEm** output was made to recruit X neurons through disinhibition driven by Ω neurons, as indicated in **Figures 5A and B₂**.

The control of LTP at synapses on inhibitors by principal cell targets is exerted via recurrent inputs to the inhibitors from their targets (**Figure 5A**). These inputs have ligand-gated synaptic effects that are somewhat parallel to those causing LTP and LTD of afferents to cerebellar Purkinje cells, which are by far the best studied of inhibitory neurons with plastic inputs. Recurrent input is taken to produce LTP by mechanisms similar to those which foster LTP at afferent synapses on Purkinje cells. There, moderate input is thought to initiate LTP by causing modest Ca^{2+} elevations (see Lev-Ram et al., 2003; Coesmans et al., 2004; Jorntell and Hansel, 2006). Since extinction in FRAT is mostly due to growth of inhibition rather than erasure of potentiation of principal cell synapses, reconditioning of a previously extinguished FRAT is achieved by erasing the inhibition that was established by extinction training. This is effected by strong input to inhibitors from R-cells. Such strong input causes large “Ca” elevations, which in the presence of reinforcing neuromodulator, initiates LTD of synapses made by active afferents on the inhibitors. This is analogous to the LTD of Purkinje cells synapses produced by climbing fiber input.

Quantification of postsynaptic factors affecting change. The graphs of **Figure 5C** quantify the effects and interactions of ordinary and neuromodulatory input to principal cell and inhibitory interneurons, as described in the preceding paragraphs. In the presence of R-cell neuromodulator, active synapses on principal cells potentiate in proportion to postsynaptic depolarization, above some threshold. In the presence of X-cell neuromodulator they may undergo slight LTD, but the parameter ζ (**Figure 5C**, left), which controls the extent of such depression is set to zero in the simulations of this paper. In the presence of X-cell neuromodulator (**Figure 5C**, right), active afferent synapses on inhibitory interneurons undergo LTP to an extent determined by the amount of “Ca” that enters the interneuron due to recurrent input or input from R. The increment of LTD increases with “Ca” concentration above a rather low threshold and then rapidly saturates with increasing concentrations. In the presence of R-cell modulator, LTD develops as a function of “Ca” concentration to an extent that mirrors the relationship between increments of LTP in principal cells and depolarization.

R-cell effects always dominate those of X-cells because R inhibits X absolutely (**Figure 5A**). Extinction never develops during a protracted US because US activity also inhibits X absolutely.

Presynaptic determinants of eligibility for change. The changes in synaptic strength just described occur only at active synapses. However, activity *per se* must be neither necessary nor sufficient, because learning fails if CS onset does not actually precede US onset, and USs that occur just as CSs terminate do cause conditioning. This caused Sutton and Barto (1981) to introduce the concept of *eligibility* for change. In FRAT eligibility is established by receipt of transmitter from active CS and context-representing neurons. However, as with NMDA receptors, which

cannot open in response to postsynaptic depolarization until some milliseconds after glutamate binds to them and which continue to be responsive to depolarization for some milliseconds after unbinding (Magleby, 2004; Clarke and Johnson, 2006), *eligibility* slightly lags transmitter binding. In FRAT synapses become eligible for change one computational time step after the onset of presynaptic activity (i.e., the variable *eligibility* goes from zero toward unity) and retain eligibility for one time step after cessation of activity. Thus, onset of presynaptic activity must precede the occurrence of other change-promoting signals. Moreover, because *eligibility* is continued during the time step after CS offsets, cue conditioning is successful even if the US, instead of co-terminating with the CS, begins contiguously with CS offset. Eligibility in FRAT is a continuous variable that, above a threshold is proportional to activation of the input but that saturates at unity when activation reaches a certain level, which is a parameter of the model.

Calculation of change. The three factors of *eligibility*, R or X activity, and the effects of postsynaptic depolarization in principal (P) cells or calcium elevation in inhibitory interneurons (I) combine as the product of three corresponding numerical factors:

$$\begin{aligned} \Delta \text{Strength } P\text{-cell synapse} \\ = \begin{cases} A_R \cdot R_p(\text{Depol}) \cdot \text{Eligibility} & \text{if } A_R > 0 \\ A_X \cdot X_p(\text{Depol}) \cdot \text{Eligibility} & \text{if } A_X > 0 \end{cases} \end{aligned} \quad (3)$$

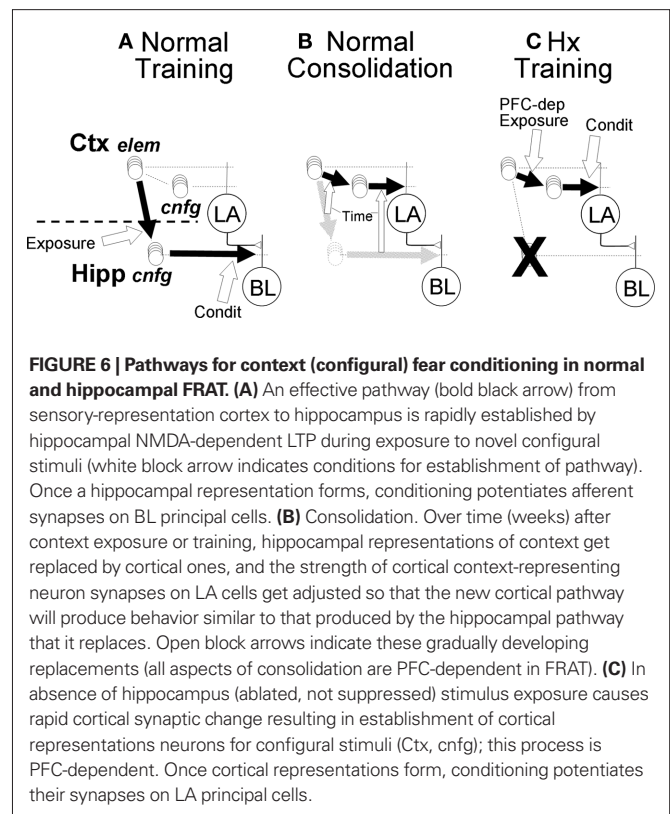
$$\begin{aligned} \Delta \text{Strength } I\text{-cell synapse} \\ = \begin{cases} A_R \cdot R_i(\text{"Ca"}) \cdot \text{Eligibility} & \text{if } A_R > 0 \\ A_X \cdot X_i(\text{"Ca"}) \cdot \text{Eligibility} & \text{if } A_X > 0 \end{cases} \end{aligned} \quad (4)$$

Eligibility is as defined above. A_R and A_X are the activations of R and X neurons, as discussed above. R_p , X_p , R_i , and X_i are the relative change factors for principal cells and inhibitors as given in the graphs of Figure 5C.

Plasticity of the non-BL LA-to-CEM pathway. If and only if BL is suppressed, coincident activity of LA principal cells and US-evoked activity of CEM neurons causes an effective pathway from BL to CEM to become established. This pathway is inhibited (but remains established) if suppression of BL is discontinued.

Forming representations of configural stimuli

Although there has been discussion in the literature as to how hippocampal circuitry might work to establish conjunctive representations (Marr, 1971; Skaggs and McNaughton, 1992; O'Reilly and McClelland, 1994), in FRAT this was simply assumed to occur. The numbers of hippocampal neurons (or in absence of hippocampus, of cortical conjunction-representing neurons) that can be activated by a stimulus and their levels of activation grow with negative acceleration whenever the stimulus is present (see Figures 6A and 10). About 20 s of exposure are needed before FRAT hippocampal representations develop to a point where they can have any impact on behavior, and it requires several minutes of exposure before activity is maximal. In the absence of hippocampus (Figure 6C), cortical representations grow similarly but a little slower (see simulations



of Figure 10, which are roughly consistent with experimental observation Wiltgen et al., 2006). Once established, hippocampal representations are projected to BL, and cortical representations to LA, to provide bases for excitatory and inhibitory conditioning as described above.

Consolidation requires the gradual and presumably experience-independent transfer of hippocampal conjunctive representations to representation cortex. Since representation cortex innervates LA rather than BL, it is also necessary to establish new synaptic weights that allow the newly functional circuitry to mimic that which it replaces (Figure 6B). Algorithms for doing this biologically have not been proposed, but the process is believed to depend on off-line neural activity, perhaps during sleep (Marr, 1971; Wilson and McNaughton, 1994; McClelland et al., 1995; Ji and Wilson, 2007). In FRAT the replacement was done one BL input channel (A, B, C, A1, A2, etc.) at a time: First the CEM output produced by a controlled BL input with inhibition disabled was determined and the synaptic weight on LA principal cells that would be needed in order for a similar LA input to produce the same output was calculated. Then BL principal cells were provided just enough excitation to bring CEM to its maximum firing rate, the reduction produced by BL input to the inhibitor was determined, and the synaptic weight on LA inhibitors needed to produce a similar reduction in CEM activity was calculated. Once the new weights were determined, the hippocampus–BL circuit was replaced by its cortex–LA correspondent (see Supplementary On-line Materials for details). We believe that plausible neural algorithms that involve off-line activity and synaptic plasticity could be invented.

Requirement for, and implications of, distal inhibition of P cells

Once representations of context and context-CS conjunctions have become cortical so that all input is to LA, AAB renewal cannot occur if excitation and inhibition are directed to a common electrical compartment, because the switch to context B removes comparable amounts of inhibition and excitation. However, if inhibition is proximal and G_c is large, the value of the left bracket of Eq. 2 decreases little, while that of the right bracket increases substantially, and renewal occurs. This is why we have used two-compartment models of principal cell dendrites and directed extinction-causing inhibition to the proximal compartment.

Implementation of experimental procedures

In the next section the freezing responses of FRAT to simulations of various ablations, inactivations, and pharmacological manipulations are discussed. The manipulations were done as follows:

Suppression of hippocampus (Hs) is emulated by reversibly setting the activations of all hippocampal neurons to zero and disabling exposure-dependent formation of hippocampal representations, thus disabling incidental learning.

Ablation of hippocampus (Hx) is emulated by additionally enabling the exposure-dependent development of cortical representations of configural stimuli, as described above. Hx is irreversible.

Suppression of LA (LAs) is emulated by reversibly setting all LA neuron responses (depolarizations, activity, etc.) to zero. Given the rules described above, this prevents all new synaptic change, but existing change is retained and can again be expressed if suppression is discontinued.

Suppression of BL (BLs) is emulated in the same manner as was LAs, but additionally the plasticity of the LA–CEm pathway (as described above) is enabled. LA and BL suppression disable the part of the consolidation algorithm that adjusts amygdala synaptic weights, but permit the establishment of cortical configural representations.

Suppression of CEm (CEMs) is emulated by reversibly setting the activity of CEm neurons to zero.

Suppression of PFC (PFCs) is emulated by reversibly disabling development of consolidated or compensatory cortical responsivity to configural cues. During the period of suppression any previously developed responses of cortical configuration-representing neurons is also attenuated. This last feature is considered especially provisional.

The effects of opiate receptor antagonism were simulated by attenuating the activity of Ω .

The effect of PAG GABA-agonist (antagonist) was simulated by attenuating (augmenting) activation of X by CEm.

BEHAVIOR OF THE MODEL

Basic properties of conditioning and extinction

The behavior of FRAT is consistent with all of the design targets of Table 1. Some of these features are illustrated here; others are illustrated in Section “Supplementary On-line Materials.”

FRAT also has a number of behavioral characteristics that are consistent with available data but that were not considered in designing the model, and still others that constitute untested predictions. Some of these are listed in Tables 3 and 4, and a few are illustrated in this section. Unless otherwise indicated, all simulated

Table 3 | Effects of hippocampal ablation (Hx) or suppression (Hs) during extinction and renewal testing on extinction (Ext) and renewal (Rnl) in FRAT (suppression continues during renewal testing).

		Hx early	Hs early	Hx late	Hs late
AAB	Ext	Ok	Slow	Ok	Ok
	Rnl	Ok	No	Ok	Ok
ABA	Ext	Ok	Slow	Ok	Slow
	Rnl	Ok	No	Ok	No
ABC	Ext	Ok	Slow	Ok	Slow
	Rnl	Ok	No	Ok	No

experiments were preceded by a session in which FRAT was given sufficient exposure to all three contexts to establish full representations for them.

Conditioning, extinction, and context. Figure 7 shows some of the basic properties of the behavior produced by FRAT.

Figures 7A–C illustrate the effects of five CS conditioning trials in context A followed by extinction of the CS in either context A or in new context B. Conditioning of fear to both the CS and the context develop rapidly and extinguish relatively slowly. As with rats, CS fear acquired in context A is also well expressed in context B.

Also, as with rats a precise evaluation of CS fear in the conditioning context is complicated by the presence of contextual fear. To provide a less contaminated evaluation of CS fear in A, Figure 7B shows a simulation in which extinction of the CS in context A was preceded by 50 min of contextual fear extinction without any CS occurrences. The Figure shows that even in the absence of background contextual fear, freezing to the CS was strong, but it then itself extinguished about as it did in Figure 7A. Also note in Figure 7B that occurrence of the CS at the start of extinction caused some second-order conditioning of fear to the context.

After extinction in any context, FRAT shows renewal of fear when the extinction context is changed. This is seen in Figure 7D for AAB, ABA, and ABC renewal. Evaluation of ABA renewal is complicated by the return of context fear when FRAT is returned to its training context (A); thus a test of renewal after extinguishing fear of A was also done. Renewal occurs with all three experimental paradigms, but it tends to be greatest for ABA, next for ABC, and least for AAB. This order seems to be consistent with available data for the behavior of rats (Bouton, 2002; Thomas et al., 2003). In the model, responsiveness during the renewal test is less for AAB than for ABC renewal because during tests in a new context (here, B or C) the only thing suppressing fear is inhibition conditioned to the CS itself. And more inhibition becomes associated to the CS during extinction in A than elsewhere because FRAT is more fearful during extinction in A (where CS, A, and the A/CS compound all promote CEm, and hence X, activity) than elsewhere (where only the CS itself promotes CEm and X activity). Responsiveness during a renewal test is greater for ABA than ABC paradigms because after extinction in B, return to A reinstates fear of both A and the A/CS compound whereas a shift to C is movement to a neutral context. Note, however, that the picture is complicated somewhat by the occurrence of secondary reinforcement during extinction, which is in fact a biological reality (Helmstetter and Fanselow, 1989).

Table 4 | Properties of FRAT that were not design explicit design targets.**CONSISTENT WITH AVAILABLE DATA**

- A. Weak cue conditioning is training-context-specific¹.
- B. Context-specificity of cue conditioning is unmasked following extinction in a novel context².
- C. Extinction is enhanced when responses are strong and diminished when responses are weak³.
- D. Effectiveness of renewal is in the order (greatest to least) ABA, ABC, AAB⁴.
- E. Rate of extinction of new cued fear is substantially reduced by post-conditioning suppression of hippocampus but not by ablation⁵.
- F. Newly acquired CS fear extinguished with hippocampus-suppressed does not show ABC or AAB renewal⁶.
- G. Suppression or ablation of hippocampus soon after extinction of newly acquired CS fear causes partial loss of extinction and a loss of extinction's context-specificity⁷.
- H. Well-conditioned CSs suppress post-shock freezing, and this suppression is opiate-dependent⁸.
- I. Fear acquired normally cannot be extinguished if BL is suppressed during extinction training⁹.

SOME UNTESTED PREDICTIONS

- J. Extended post-training suppression of BLA prevents systems-type consolidation¹⁰.
- K. Pre- or post-training PFC lesions prevent systems-type consolidation.
- L. Extinction of both cue and context fear show systems-type consolidation. After such consolidation, context shifts cause renewal, independent of hippocampus.
- M. If hippocampus is suppressed (not ablated) after completion of systems-type consolidation, CS extinction is normal in the conditioning context but slow and context-independent in novel contexts.
- N. After ablation of LA, cue conditioning remains possible, but it is conditioning context-specific.
- O. When cues predict shock onset, LA ablation increases context conditioning, but not as much as does removal of predictive cues.
- P. Inhibition of GABA in PAG promotes development of extinction.
- Q. BL lesions do not prevent context fear if hippocampus has been ablated.
- R. Joint pre-exposure to CS and conditioning context will enhance strength of one-trial CS conditioning.
- S. Inactivation of CEm during unreinforced responding prevents extinction.

References and Notes: ¹Hall and Honey (1990). ²Harris et al. (2000). ³Rescorla (2002, 2003). ⁴Bouton (2002), Thomas et al. (2003). ⁵Corcoran et al. (2005). ⁶Corcoran et al. (2005). See text ⁷Corcoran and Maren (2001). ⁸Fanselow and Bolles (1979a). ⁹Herry et al. (2008). ¹⁰Several hours of tetrodotoxin-induced silencing of BLA at any time during the first 2 days after contextual fear conditioning, but not thereafter, reduces long-term retention of contextual freezing (Sacchetti et al., 1999). While consistent with prediction J, this appears to us to be more akin to an effect on cellular- than on systems-type consolidation, because CS conditioning is affected in the same way.

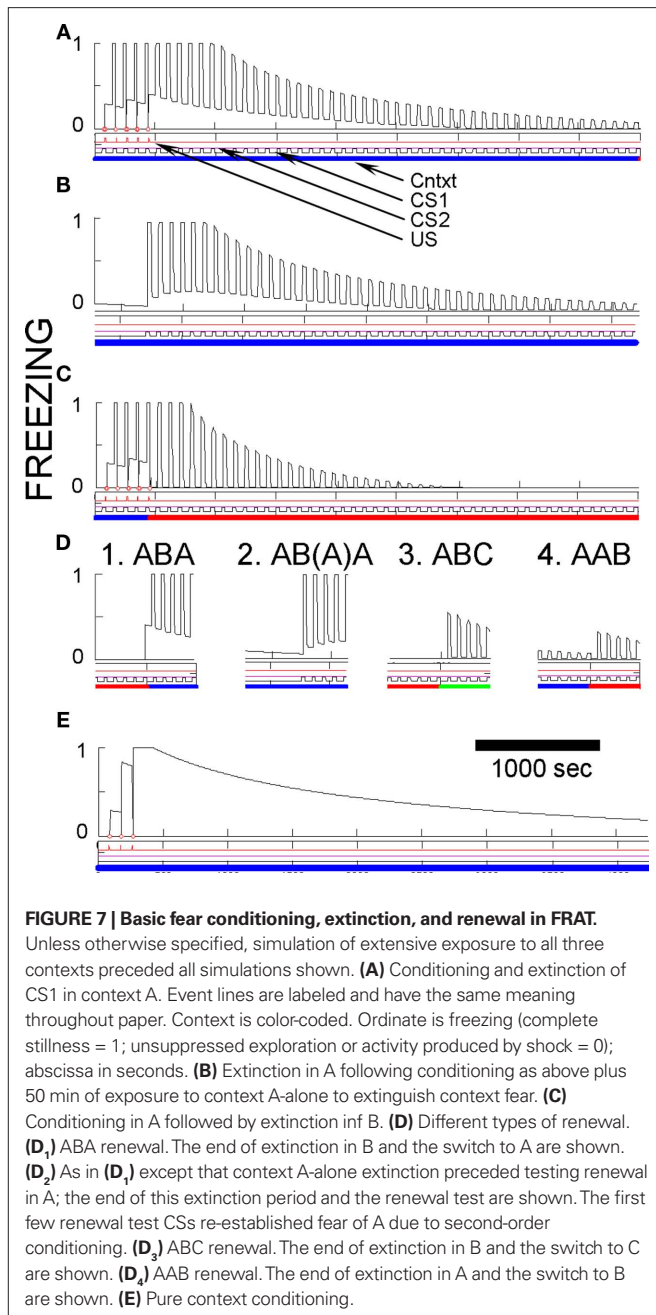
Figure 7E illustrates the effect of contextual fear training when the US is presented without any predictive CS. A much greater level of context fear develops than was seen in **Figures 7A–C**, because there is no response to the CS to drive strong opioid inhibition of US-representing neurons. Animals show a similar enhancement of context fear when there are no predictive CSs (Fanselow, 1980; Fanselow et al., 1994).

Evidence of contextual dependence of conditioning. As discussed above, cued fear established in one context seems to be fully expressed in other contexts. However, underlying fear, as reflected in the activity of CEm neurons, is considerably greater in the conditioning context than elsewhere because fear conditioned to CS/context conjunctive representations, as well as possibly behaviorally subthreshold contextual fear, are absent outside the training environment. Fear-expression in FRAT is not usually diminished away from the training context, because the dynamic range of downstream fear-expression-producing neurons is limited; very strong CEm activity in the conditioning context and less strong activity elsewhere both commonly produce maximal freezing scores. However differences can become evident if conditioning is weak. This is shown in **Figure 8**, in which training consisted of a single conditioning trial with a very weak US. A similar result is found in real rats (Hall and Honey, 1990).

Evidence of context-specificity of conditioning can also be seen in experiments where rats that had been conditioned to several CSs, each in different context, have all CSs extinguished in a new context. In such experiments (Harris et al., 2000), and in FRAT, renewal is much greater when a CS is subsequently tested in the context where it, rather than some other cue was originally conditioned (see simulations in on-line materials). Context-dependence of cued fear is also suggested by the observation that hippocampal ablation, which would be expected only to cause retrograde losses of responses conditioned to conjunctive cues, in fact causes losses in cue conditioning if the conditioning was weak (Quinn et al., 2008b); this is what would be predicted if some of the conditioning were to cue/context stimulus compounds.

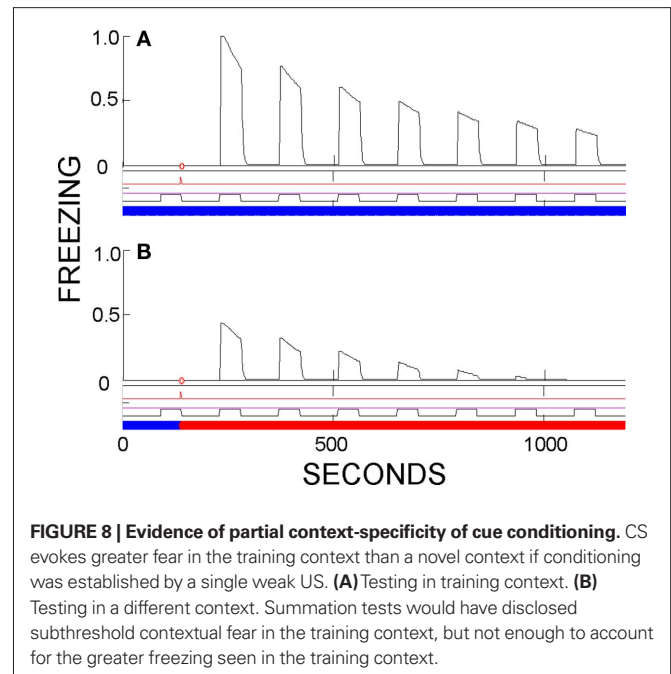
Aspects of FRAT hippocampal function

As discussed above, at least at initial learning, the hippocampus is the FRAT amygdala's primary source of information about configural stimuli, and in particular of contextual information. The hippocampal representations of configural stimuli that project to amygdala are formed fairly rapidly during exposure to the stimuli, and once these are formed, they provide input to BL where conditioning of fear to them, as well as conditioning of inhibitory activity that contributes to context-specific extinction, become possible. From this alone it would be expected that (a). Contextual fear conditioning would be entirely due to potentiation of synapses within



BL, which is where FRAT hippocampus projects. (b). Contextual fear could not be established until time has been allowed for the formation of a hippocampal representation of the context (c). Context fear conditioning should be prevented by hippocampal ablation (anterograde amnesia). (d). Post-training hippocampal removal should abolish previously learned contextual fear (retrograde amnesia). (e). The context-specificity of extinction and possibly extinction itself should be dependent on hippocampus.

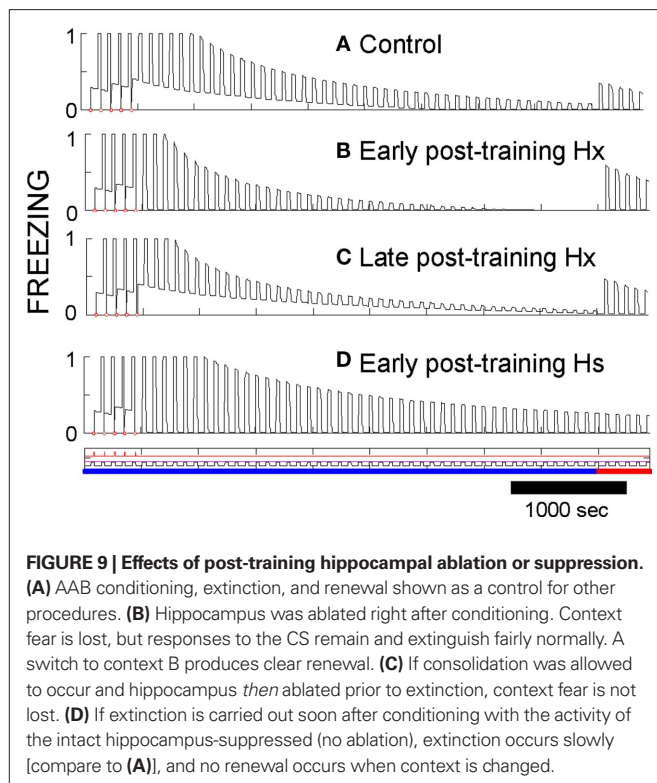
However, there are several (empirically motivated) aspects of FRATs construction that complicate this picture: (1). If hippocampus is ablated, a PFC-dependent compensation process of unspecified mechanism allows sensory-representation cortex to form configural



representations, though the formation of configural representations by this process is a little slower than normal. (2). If hippocampal plasticity or activity are suppressed pharmacologically but hippocampus remains intact, the PFC-dependent compensation process does not occur. (3). Over extended periods of time, systems-type consolidation occurs: hippocampal representations are replaced by ones located in sensory-representation cortex, and if hippocampus-to-BL principal cell or inhibitory interneuron synapses had become potentiated, potentiation that has similar consequences for downstream activity becomes established at the LA targets of the new cortical representations. In the following sections we discuss some of the behavioral consequences of these considerations.

Post-training hippocampal ablation (Hx) or suppression (Hs). Ablation of hippocampus soon after conditioning in FRAT has little effect on CS-evoked freezing but abolishes freezing to context (compare **Figures 9A,B**). If time is allowed for consolidation before ablation, context fear survives (**Figure 9C**). This phenomenology replicates standard empirical findings.

Extinction occurs in a normal manner after early post-training ablation (**Figure 9B**), but it is slow if hippocampus is suppressed rather than removed (**Figure 9D**). The slowing of extinction in hippocampus-suppressed FRAT was an unanticipated consequence of programming FRAT to meet the conditions of **Table 1**. It occurs because hippocampal representations of the context and context/CS conjunctions are missing with hippocampus-suppressed, so the only representations present during extinction are those of cortex. Consequently, less total inhibition per trial is acquired than with hippocampus intact (with ablation, as opposed to suppression, cortical representations of the extinction context and of context/CS conjunctions are rapidly established at the start of extinction). In fact, Corcoran and Maren (Corcoran et al., 2005) have observed slow extinction in rats under just these circumstances.



After extinction of hippocampectomized FRAT, renewal also occurs normally (Figures 9B,C). This is the case because cortical conjunctive representations to which inhibition gets conditioned develop rapidly at the start of extinction. However, if hippocampus is suppressed rather than ablated, AAB and ABC fail though ABA renewal persists. In rats all three types of renewal have been reported to fail under similar circumstances (Corcoran et al., 2005). This may represent a discrepancy between FRAT and reality; however, we consider it is possible that apparent lack of ABA renewal in the rats was due to ceiling effects imposed by the measurement of CS fear on a high contextual fear baseline in the experiments reported.

Table 3 summarizes a number of effects of post-conditioning, hippocampal ablations and suppressions in FRAT. It should be noted that if consolidation has been allowed to occur, extinction in the conditioning context during hippocampal suppression is normal and renewal occurs if FRAT is then tested in a different context (i.e., AAB extinction and renewal are normal). However, if extinction with hippocampus-suppressed occurs elsewhere, this is not the case. This is because consolidation establishes representations of the conditioning context and context/CS compound in cortex; but such representations for different contexts are lacking and cannot be established during extinction if hippocampus is suppressed. These would be interesting predictions to test.

If FRAT hippocampus is either ablated or suppressed soon after extinction is complete, there is partial recovery from extinction, and responsiveness becomes independent of both conditioning and extinction context (not illustrated). This occurs because all of the conjunctive representations to which inhibition has become conditioned become unavailable. Exactly this outcome was seen when rats were given hippocampal muscimol injections soon after

extinction (Corcoran and Maren, 2001) although results of a later study gave somewhat different results (Corcoran and Maren, 2004); FRAT simulates the findings of both reports if one assumes that suppression of hippocampal activity by muscimol was not quite complete (see Supplementary On-line Materials). If post-extinction hippocampal ablation or suppression are performed after FRAT consolidation mechanisms have been allowed to operate, conjunctive representations have moved to cortex and synaptic potentiation has occurred that replaces the hippocampal pathway to principal cell inhibitors (as in Figure 6A). Thus loss of hippocampus causes no loss of extinction, and renewal remains normal. We are not aware of comparable animal experiments.

The immediate shock deficit. The immediate shock deficit as seen in FRAT is illustrated in Figure 10. In the normal case shocks given after less than about 13 s of exposure to an unfamiliar context cause development of no manifest context fear, but fear develops if there has been more cumulative exposure prior to the shock. Freezing has become nearly asymptotic by around 800 s of exposure.

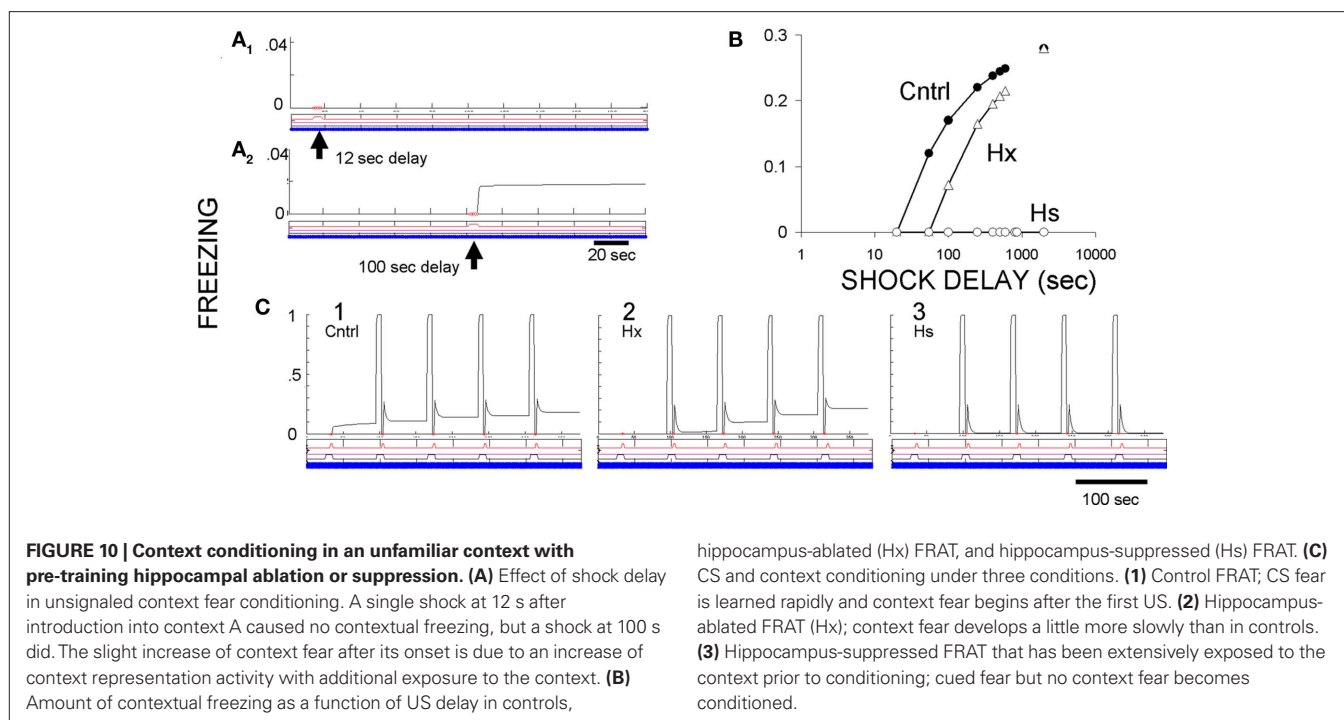
Pre-training hippocampal ablation (Hx) or suppression (Hs). If FRAT hippocampus is suppressed prior to exposure to an unfamiliar context, no fear can be conditioned to that context (Figure 10B). However, cue conditioning appears to occur normally (Figure 10C), though careful evaluation of strength of conditioning, context-dependence of conditioning, etc., discloses abnormalities (not illustrated). These characteristics follow from a total absence of any contextual input to amygdala if hippocampus has been suppressed prior to any stimulus exposure. If hippocampus is ablated rather than suppressed, compensation processes allow the formation of configural representations in cortex, so the properties of both cue and context conditioning and extinction appear fairly normal, though quantitative differences can be detected. In particular, a little more context exposure is needed prior to conditioning to achieve a given level of context fear (Figure 10B); in this, FRAT is very similar (by design) to real animals (Wiltgen et al., 2006).

Blocking, extinction, and PAG opiate receptors

As discussed above and as prescribed in the design-criteria of Table 1, PAG opiate receptor inhibitors attenuate both blocking (Figure 11A) and extinction (Figure 11B) in FRAT, as in rats (McNally et al., 2004a,b; Cole and McNally, 2007).

The discovery in rats of attenuation of extinction by opiate inhibitors and enhancement of extinction by an opiate enhancing agent (McNally et al., 2004b; McNally, 2005) would seem to imply that extinction is dependent on some signal that is itself PAG-dependent-in-effect an “extinction” signal that is extrinsic to the amygdala.

There are a number of minor variations of the circuitry controlling R and X in FRAT that would have the same ultimate effects on behavior. However, in the circuit actually employed (Figure 11) opiate antagonists, while attenuating blocking for primary reinforcement, would not do so for secondary reinforcement. Moreover, PAG GABA inhibitors would facilitate extinction and interfere with conditioning by causing continual firing of the X-cells. It would be interesting if such effects were found in animals.



DISCUSSION

The goal of this project was to try to create a neurally plausible model that reproduces a number of known empirical features of fear learning (Table 1). If the neural plausibility of the model is accepted, the goal has been met. Moreover, the neural circuitry created to meet the criteria of Table 1 generated a number of additional behavioral properties which appear to be empirically valid (Table 4; simulations in Supplementary On-line Materials). For example, the model predicts that renewal should be only partial and that the procedure ABA should be most and AAB least effective (Table 4, D); this conforms with common findings (Bouton, 2002; Thomas et al., 2003). The model also correctly predicts that suppression of hippocampus *during* extinction should slow extinction and prevent what extinction occurs from being context-dependent, whereas suppression *following* extinction, should itself cause partial, context-independent recovery from extinction (Table 4, E–G; Corcoran et al., 2005).

The model additionally makes many untested predictions, some of which are listed at the bottom of Table 4. For example, it predicts that suppression of BLA, or CEM during the consolidation period over which contextual fear becomes hippocampus-independent should prevent such consolidation (Table 4, J). It also predicts that extinction should be prevented by CEM suppression during extinction training (Table 4, S).

ASSUMPTIONS

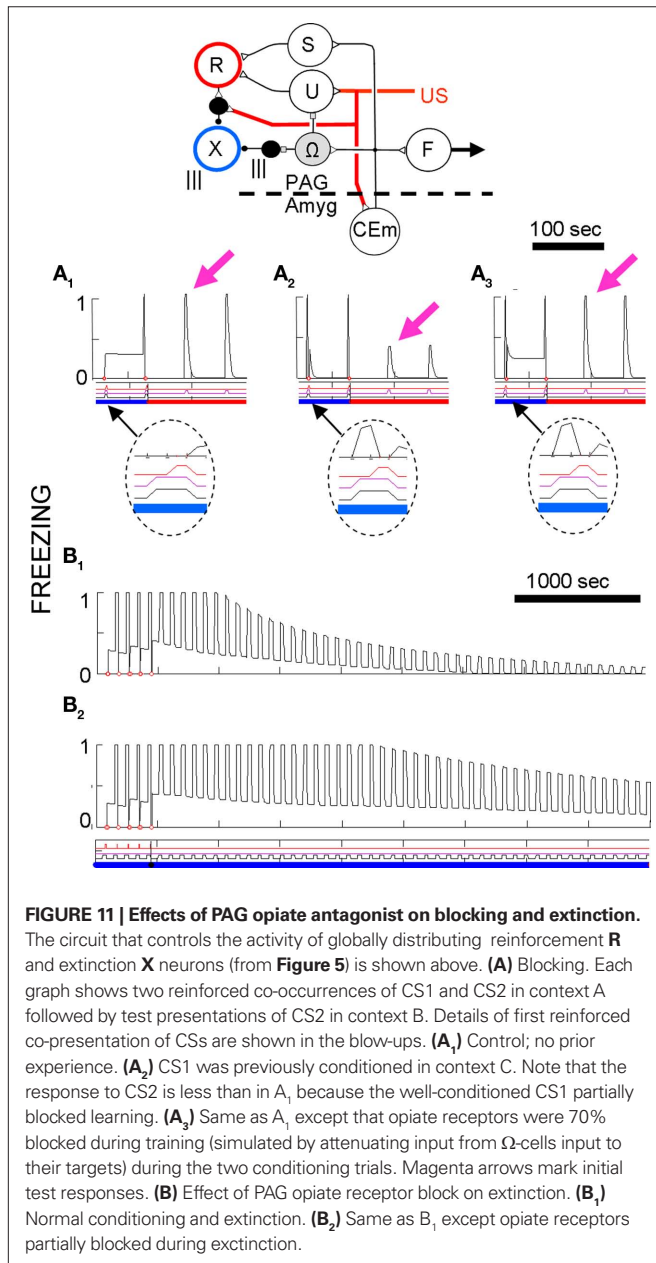
Two aspects of the model especially deserve further discussion.

Globally distributing neuromodulators

Control, of plasticity by R and X-cell neuromodulators was used in two ways in FRAT: (i). To make synaptic change contingent on the ultimate output of the amygdala circuit (as assessed by CEM

activity). This was crucial to the handling of both secondary reinforcement and blocking. Reliable renewal also depended on control of extinction by the circuit's output. In the absence of such control, inhibition at the level of LA could grow so strong during extinction training that not enough (context-independent) input from LA to BL would remain after extinction to renew fear when context shifts lead to reduced inhibition at the level of BL. (ii). To make conditioning contingent specifically on onset of USs (or in the case of secondary reinforcement, of well-trained CSs). This was crucial to prevent a new CS that comes on during a protracted US or established CS from becoming conditioned.

These matters could perhaps have been dealt with without the use of neuromodulators. However, changes in synaptic strength must in one way or another be made subject to control by the current output of the circuit, by the level of US activity, and by *increments* in each of these quantities. It is possible that increments could be computed by properties of the target cells themselves, but in FRAT the computation was implicit in the properties of the R-cells. In order for external plasticity-controlling information to be utilized, it must be represented within plastic cells in a way that does not interfere with other sorts of processing. It would seem to be difficult to mediate such control via target cell membrane potential without unwanted interactions between normal processing and plasticity control signals. Indeed, prevention of unwanted secondary reinforcement by CS-caused depolarization of principal cells was the immediate impetus to positing a neuromodulatory component to the R-cell signal in FRAT. It might have been possible to avoid this, but designing the circuit was made easier by assuming that external plasticity-controlling signal work via the cell's internal chemical signaling systems, i.e., that they are neuromodulatory.



Neuromodulatory control of conditioning of course has precedent in VTA dopamine system-dependent learning (Schultz and Dickinson, 2000). Fear learning or BLA plasticity also seems to be controlled, at least in some circumstances, by dopamine and/or norepinephrine (Selden et al., 1991; Johnston et al., 1999; Bissiere et al., 2003; Pezze and Feldon, 2004; Tully et al., 2007; Fadok et al., 2009, 2010; Bush et al., 2010; Iordanova, 2010; Johansen et al., 2010a).

The presumption that an external signal influences extinction has somewhat less precedent. However, discussions of the midbrain dopamine system often imagine a role for negative error signals in extinction, and there is evidence for the importance of such signals in extinction of cerebellar eye-blink conditioning (Medina et al., 2002). The discovery that extinction of fear conditioning in rats

can be attenuated by opiate inhibitors and enhanced by an opiate enhancing agent (McNally et al., 2004b; McNally, 2005) would seem to imply that extinction is dependent on some signal that is itself PAG-dependent, and thus dependent on events external to the amygdala.

It is known that dopamine neurons, whose increased activity may mediate reinforcement, slow their ongoing rates of firing when expected reinforcements are omitted and that extinction of cerebellar eye-blink responses is at least in part due to slowed spontaneous firing of climbing fibers when unreinforced eye-blink responses occur (Medina et al., 2002). This suggests a single external signal whose increase promotes reinforcement and decrease promotes extinction. However, such bipolar control requires neurons with substantial baseline firing rates, and we elected instead to use separate reinforcing and extinction-promoting signals, carried by the **R** and the **X**-cells, respectively. One consequence of this, given the effects of opiate transmitter system alterations, is that **CEm** had to recruit **X** by disinhibition, whereas if the global extinction signal had been reduction in spontaneous **R** firing, this complexity would have been avoided. The matter is probably subject to experimental test: If extinction-promoting signals are actually recruited via disinhibition, then it should be possible to prevent extinction with PAG GABA-agonists, and to augment extinction by use of antagonists (see Supplementary On-line Materials).

Heavy use of context/cue conjunctive representations

Learned representations of context/cue conjunctions are crucial to the return of extinguished CS fear when FRAT is moved from contexts in which extinction was carried out (i.e., renewal). The removal of inhibition conditioned to the context alone can cause some renewal, but in real rats renewal of responding to a CS is at least partially dependent on removing the animal from the particular context in which that particular CS was extinguished (e.g., Harris et al., 2000). Related observations are what led Bouton to suggest that the CS-driven inhibition postulated to mediate extinction of a CS is “gated” by context (Bouton and Ricker, 1994), and it is what caused us to introduce context/cue conjunctive representations in FRAT. In FRAT these representations are established in hippocampus. But as pointed out below, it might also be possible to have them arise by convergence on BL cells of CS information from LA and context information from hippocampus.

Given other assumptions, the existence of context/cue conjunctive representations and the requirement that they develop within hippocampus have several implications, some of which have been partially verified: (1). If cue conditioning is weak, cued fear should be greater in the context of conditioning (where conditioned context/cue representations are present) than elsewhere (see Hall and Honey, 1990; Harris et al., 2000; and **Figure 8**). (2). After LA lesions, cued fear learning to context/CS conjunctive representations, which are effective in BL, should still be possible, but such cued fear should be manifest only in the context of conditioning (**Table 4, M**). (3). Post-training hippocampal ablation should cause some retrograde amnesia for weak cued fear if there is a degree of overlap between training and test contexts (see Quinn et al., 2008b). (4). Manifestations of conditioning to context/cue com-

pounds should require hippocampus-dependent pre-conditioning experience with the to-be-conditioned CS in the context where conditioning will occur (Table 4, R).

UNFINISHED BUSINESS

Necessity of model properties

Our goal in constructing FRAT was to create an “automaton” that would emulate properties of biological fear conditioning in a biologically plausible way, but biological plausibility and parsimony sometimes conflicted. For example, LTP on inhibitory interneurons must not be allowed to go too far, or renewal will be liable to fail. Two mechanisms served to limit such LTP: (1) X-cells were made to have a sufficiently high threshold so that the neuromodulator required for inhibitory neuron synapse LTP is not released once CEm activity falls below a certain point (see Figure 5B₂). (2) LTP at inhibitor synapses cannot occur if “Ca” level produced in inhibitors by recurrent input from principal cells falls too low (see Figure 5C, right). It may well be that the first mechanism, which is probably the more important, is sufficient to ensure failures of renewal. But the second mechanism conforms to the biological reality that LTP and LTD commonly depend on postsynaptic Ca²⁺ levels. Of course there is undoubtedly considerable redundancy in biological systems. Nevertheless, one important piece of unfinished business is to determine how important individual aspects of the model are to successful simulation of various attributes of fear conditioning.

Establishment and consolidation of configural representations and their effects

There were several points during the construction of FRAT where problems were solved by fiat without creating neural circuitry that could accomplish the needed computations. Thus: (i) Hippocampal representations of configural cues were assumed to form at defined rates during stimulus exposure. (ii) A slightly slower process with similar outcome was assumed to lead to creation of cortical conjunctive representations if hippocampus was ablated, and the formation of these cortical representations was PFC-dependent. No mechanisms for either process were postulated, nor was any mechanism for involving PFC in the latter process suggested. Nor was any mechanism postulated for detecting the absence, as opposed to the silence of the hippocampus after its ablation. (iii) Configural cue representations that had been formed in hippocampus were assumed to gradually shift to representation cortex, and their projections to LA were adjusted so that they could functionally replace the hippocampal–BL connections they replaced (consolidation). Although a detailed algorithm for accomplishing this was postulated, no neural mechanisms for implementing the algorithm was offered. (iv) Utilization of cortical representations of configural cues was made somewhat PFC-dependent, but no theory was offered as to why this might be so.

There has been considerable thinking about the formation of hippocampal representations (Marr, 1971; Skaggs and McNaughton, 1992; O’Reilly and McClelland, 1994), and this could provide guidance in creating a model of that process, but, as far as we know, the other matters have not previously been considered at a theoretical or computational level. Ever since Marr (1971) first presented his theory of hippocampal function, it has been conjectured that con-

solidation requires neural activity. Although experimental evidence for a role of off-line activity associated with consolidation has been sought (Marr, 1971; Wilson and McNaughton, 1994; McClelland et al., 1995; Ji and Wilson, 2007), we know of no concrete algorithms using such activity that have been proposed. In FRAT we found it necessary to set the weights of consolidated pathway synapses so that the output of the amygdala circuit in response to input under various conditions would be the same before and after consolidation. Thus, in-effect we found it necessary to use an activity-dependent mechanism. Neurally plausible algorithms for accomplishing this are needed.

The representation of time

Some kinds of conditional responses are well-timed events, with the CR beginning slightly before the expected onset time of the US, and there have been theoretical attempts to deal with this, such as the concept of the serial compound stimulus (Sutton and Barto, 1990) used in discussions of primate VTA dopamine neuron responses (Schultz et al., 1997). Time has been less a focus in fear conditioning studies, and the representation of time plays no role in FRAT. In models where CS duration is explicitly coded, it is possible to condition expectation of the US at a particular time after CS onset, and the failure of the US to occur at the expected time during extinction can generate an error signal that helps trigger changes responsible for extinction. In FRAT, which lacks such a signal, incremental changes responsible for extinction simply occur in each time interval during which there is fear but no reinforcement. A consequence of this is that extinction rate is affected rather strongly by CS duration during extinction; short CSs cause much less extinction per trial than longer ones (not illustrated). There is some evidence that this prediction is empirically valid (Shipley, 1974).

Heterogeneous responses of BLA neurons

We made the initial simplifying assumption that LA and BL cells of FRAT are uniform populations of neurons, each innervated by all the afferents to the region via plastic synapses, and that learning and extinction of fear responses do not involve plasticity of synapses downstream of BLA. However, if there were such uniform innervation of LA and BL neurons, and if, as is likely, stimuli are coded by groups of neurons each of which also contributes to the representation of other stimuli, then after conditioning to numerous stimuli, many synapses of the circuit would be potentiated, and over-generalization to novel stimuli would be likely.

This problem can probably be greatly reduced by abandoning the assumption of uniform innervation and assuming that each BLA principal cell is innervated by different, though partially overlapping populations of afferents, or that competitive mechanisms of the kind recently discussed by Silva and colleagues (Zhou et al., 2009) force different BLA cells to mediate responsiveness to different stimuli. Under such circumstances downstream plasticity should help to make conditioning stimulus-specific (see, e.g., Brindley, 1969). It is in fact the case that plasticity downstream of BLA has been found to be important for both learning and extinction (Wilensky et al., 2006; Amano et al., 2010).

If the cells of BLA are heterogeneously innervated, then different cells of BL will be innervated by different combinations of LA neurons and by different contextual representation neuron from hippocampus. This alone should lead to at least some degree of context-specificity of extinction even without the use of context/CS conjunctive input from hippocampus. A related approach was taken in a recent computational model of BL (Vlachos et al., 2011) aimed at explaining the “extinction” neurons found there (Herry et al., 2008). Whether context/CS conjunctive representations that develop within hippocampus would in this case still be needed to produce realistic renewal behavior will depend on quantitative considerations that cannot be pre-judged. However, whether essential or not, it seems plausible that context/CS conjunctive representations should in fact form in hippocampus and should be among the representations to which fear becomes conditioned.

Unsimulated phenomena

Although FRAT simulates many known attributes of fear conditioning (Tables 1 and 4), there are others that the model as it now stands does not simulate. Some of these could be incorporated by small elaborations of features already present. For example,

REFERENCES

- Albert, M., and Ayres, J. J. B. (1997). One-trial simultaneous and backward excitatory conditioning in rats: lick suppression, freezing, and rearing to CS compounds and their elements. *Anim. Learn. Behav.* 25, 210–220.
- Amano, T., Unal, C. T., and Pare, D. (2010). Synaptic correlates of fear extinction in the amygdala. *Nat. Neurosci.* 13, 489–494.
- Amorapanth, P., Nader, K., and Ledoux, J. E. (1999). Lesions of periaqueductal gray dissociate-conditioned freezing from conditioned suppression behavior in rats. *Learn. Mem.* 6, 491–499.
- Anagnostaras, S. G., Gale, G. D., and Fanselow, M. S. (2001). Hippocampus and contextual fear conditioning: recent controversies and advances. *Hippocampus* 11, 8–17.
- Anglada-Figueroa, D., and Quirk, G. J. (2005). Lesions of the basal amygdala block expression of conditioned fear but not extinction. *J. Neurosci.* 25, 9680–9685.
- Ayres, J. J. B., Haddad, C., and Albert, M. (1987). One-trial excitatory backward conditioning as assessed by conditioned suppression of licking in rats: concurrent observations of lick suppression and defensive behaviors. *Anim. Learn. Behav.* 15, 212–217.
- Bast, T., Zhang, W. N., and Feldon, J. (2003). Dorsal hippocampus and classical fear conditioning to tone and context in rats: effects of local NMDA-receptor blockade and stimulation. *Hippocampus* 13, 657–675.
- Bauer, E. P., and LeDoux, J. E. (2004). Heterosynaptic long-term potentiation of inhibitory interneurons in the lateral amygdala. *J. Neurosci.* 24, 9507–9512.
- Bauer, E. P., Schafe, G. E., and Ledoux, J. E. (2002). NMDA receptors and L-type voltage-gated calcium channels contribute to long-term potentiation and different components of fear memory formation in the lateral amygdala. *J. Neurosci.* 22, 5239–5249.
- Biedenkapp, J. C., and Rudy, J. W. (2009). Hippocampal and extrahippocampal systems compete for control of contextual fear: role of ventral subiculum and amygdala. *Learn. Mem.* 16, 38–45.
- Bissiere, S., Humeau, Y., and Luthi, A. (2003). Dopamine gates LTP induction in lateral amygdala by suppressing feedforward inhibition. *Nat. Neurosci.* 6, 587–592.
- Blair, H. T., Schafe, G. E., Bauer, E. P., Rodrigues, S. M., and Ledoux, J. E. (2001). Synaptic plasticity in the lateral amygdala: a cellular hypothesis of fear conditioning. *Learn. Mem.* 8, 229–242.
- Blanchard, R. J., Fukunaga, K. K., and Blanchard, D. C. (1976). Environmental control of defensive reactions to footshock. *Bull. Psychon. Soc.* 8, 129–130.
- Bolles, R. C., and Fanselow, M. S. (1980). A perceptual-defensive-recuperative model of fear and pain. *Behav. Brain Sci.* 3, 291–301.
- Bouton, M., and Ricker, S. (1994). Renewal of extinguished responding in a second context. *Anim. Learn. Behav.* 22, 317–324.
- Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychol. Bull.* 114, 80–99.
- Bouton, M. E. (2002). Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. *Biol. Psychiatry* 52, 976–986.
- Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learn. Mem.* 11, 485–494.
- Bouton, M. E., and King, D. A. (1983). Contextual control of the extinction of conditioned fear: tests for the associative value of the context. *J. Exp. Psychol. Anim. Behav. Process.* 9, 248–265.
- Brindley, G. S. (1969). Nerve net models of plausible size that perform many simple learning tasks. *Proc. R. Soc. Lond. B Biol. Sci.* 174, 173–191.
- Brunzell, D. H., and Kim, J. J. (2001). Fear conditioning to tone, but not to context, is attenuated by lesions of the insular cortex and posterior extension of the intralaminar complex in rats. *Behav. Neurosci.* 115, 365–375.
- Burgos-Robles, A., Vidal-Gonzalez, I., Santini, E., and Quirk, G. J. (2007). Consolidation of fear extinction requires NMDA receptor-dependent bursting in the ventromedial prefrontal cortex. *Neuron* 53, 871–880.
- Bush, D. E., Caparosa, E. M., Gekker, A., and Ledoux, J. (2010). Beta-adrenergic receptors in the lateral nucleus of the amygdala contribute to the acquisition but not the consolidation of auditory fear conditioning. *Front. Behav. Neurosci.* 4:154. doi: 10.3389/fnbeh.2010.00154
- Calandrea, L., Desmedt, A., Decorte, L., and Jaffard, R. (2005). A different recruitment of the lateral and basolateral amygdala promotes contextual or elemental conditioned association in Pavlovian fear conditioning. *Learn. Mem.* 12, 383–388.
- Carrière, P. (1993). The periaqueductal gray and defensive behavior: functional representation and neuronal organization. *Behav. Brain Res.* 58, 27–47.
- Chapman, P. F., Kairiss, E. W., Keenan, C. L., and Brown, T. H. (1990). Long-term synaptic potentiation in the amygdala. *Synapse* 6, 271–278.
- Cho, Y. H., Friedman, E., and Silva, A. J. (1999). Ibotenate lesions of the hippocampus impair spatial learning but not contextual fear conditioning in mice. *Behav. Brain Res.* 98, 77–87.
- Ciocchi, S., Herry, C., Grenier, F., Wolff, S. B., Letzkus, J. J., Vlachos, I., Ehrlich, I., Sprengel, R., Deisseroth, K., Stadler, M. B., Müller, C., and Luthi, A. (2010). Encoding of conditioned fear in central amygdala inhibitory circuits. *Nature* 468, 277–282.
- Clarke, R. J., and Johnson, J. W. (2006). NMDA receptor NR2 subunit dependence of the slow component of magnesium unblock. *J. Neurosci.* 26, 5825–5834.
- Coesmans, M., Weber, J. T., De Zeeuw, C. I., and Hansel, C. (2004). Bidirectional parallel fiber plasticity in the cerebellum under climbing fiber control. *Neuron* 44, 691–700.
- Cole, S., and McNally, G. P. (2007). Opioid receptors mediate direct predictive

spontaneous recovery would automatically emerge if passage of time were made to alter context (Bouton, 1993). Other phenomena would require somewhat greater additions. Of particular importance in this regard are a number of effects that are thought to be due to shifts of attention to or from particular CSs as a result of past experience (McNally et al., 2004a; Iordanova et al., 2006; McNally and Westbrook, 2006). Such phenomena have motivated or supported major psychological theories of conditioning such as the Mackintosh or Pearce–Hall models (Mackintosh, 1975; Pearce and Hall, 1980). The addition of attentional mechanisms is thus an important topic for development.

CONCLUSION

It should be clear from the present paper that computational models of the kind employed here can help to make sense of findings that previously lacked explanation and can also be a rich source of untested hypotheses. It is also clear from the preceding discussion that there is considerable scope for future work.

ACKNOWLEDGMENTS

This work was supported in part by NIH Grant R01 MH602122 to Michael S. Fanselow.

- fear learning: evidence from one-trial blocking. *Learn. Mem.* 14, 229–235.
- Corcoran, K. A., Desmond, T. J., Frey, K. A., and Maren, S. (2005). Hippocampal inactivation disrupts the acquisition and contextual encoding of fear extinction. *J. Neurosci.* 25, 8978–8987.
- Corcoran, K. A., and Maren, S. (2001). Hippocampal inactivation disrupts contextual retrieval of fear memory after extinction. *J. Neurosci.* 21, 1720–1726.
- Corcoran, K. A., and Maren, S. (2004). Factors regulating the effects of hippocampal inactivation on renewal of conditional fear after extinction. *Learn. Mem.* 11, 598–603.
- Di Scala, G., Mana, M. J., Jacobs, W. J., and Phillips, A. G. (1987). Evidence of Pavlovian conditioned fear following electrical stimulation of the periaqueductal grey in the rat. *Physiol. Behav.* 40, 55–63.
- Doron, N. N., and Ledoux, J. E. (1999). Organization of projections to the lateral amygdala from auditory and visual areas of the thalamus in the rat. *J. Comp. Neurol.* 412, 383–409.
- Esmoris-Arranz, F. J., Pardo-Vazquez, J. L., and Vazquez-García, G. A. (2003). Differential effects of forward or simultaneous conditioned stimulus-unconditioned stimulus intervals on the defensive behavior system of the Norway rat (*Rattus norvegicus*). *J. Exp. Psychol. Anim. Behav. Process.* 29, 334–340.
- Fadok, J. P., Darvas, M., Dickerson, T. M., and Palmiter, R. D. (2010). Long-term memory for Pavlovian fear conditioning requires dopamine in the nucleus accumbens and basolateral amygdala. *PLoS ONE* 5, e12751. doi: 10.1371/journal.pone.0012751
- Fadok, J. P., Dickerson, T. M., and Palmiter, R. D. (2009). Dopamine is necessary for cue-dependent fear conditioning. *J. Neurosci.* 29, 11089–11097.
- Fanselow, M. S. (1980). Signaled shock-free periods and preference for signaled shock. *J. Exp. Psychol. Anim. Behav. Proc.* 6, 65–80.
- Fanselow, M. S. (1986). Associative vs topographical accounts of the immediate shock-freezing deficit in rats: implications for the response selection rules governing species-specific defensive reactions. *Learn. Motiv.* 17, 16–39.
- Fanselow, M. S. (1990). Factors governing one-trial contextual conditioning. *Anim. Learn. Behav.* 18, 264–270.
- Fanselow, M. S. (1998). Pavlovian conditioning, negative feedback, and blocking: mechanisms that regulate association formation. *Neuron* 20, 625–627.
- Fanselow, M. S. (2000). Contextual fear, gestalt memories, and the hippocampus. *Behav. Brain Res.* 110, 73–81.
- Fanselow, M. S. (2010). From contextual fear to a dynamic view of memory systems. *Trends Cogn. Sci. (Regul. Ed.)* 14, 7–15.
- Fanselow, M. S., and Bolles, R. C. (1979a). Naloxone and shock-elicited freezing in the rat. *J. Comp. Physiol. Psychol.* 93, 736–744.
- Fanselow, M. S., and Bolles, R. C. (1979b). Triggering of the endorphin analgesic reaction by a cue previously associated with shock: reversal by naloxone. *Bull. Psychon. Soc.* 14, 88–90.
- Fanselow, M. S., and Dong, H. W. (2010). Are the dorsal and ventral hippocampus functionally distinct structures? *Neuron* 65, 7–19.
- Fanselow, M. S., and Kim, J. J. (1994). Acquisition of contextual Pavlovian fear conditioning is blocked by application of an NMDA receptor antagonist D,L-2-amino-5-phosphonovaleric acid to the basolateral amygdala. *Behav. Neurosci.* 108, 210–212.
- Fanselow, M. S., Kim, J. J., Yipp, J., and De Oca, B. (1994). Differential effects of the N-methyl-D-aspartate antagonist DL-2-amino-5-phosphonovalerate on acquisition of fear of auditory and contextual cues. *Behav. Neurosci.* 108, 235–240.
- Frankland, P. W., and Bontempi, B. (2005). The organization of recent and remote memories. *Nat. Rev. Neurosci.* 6, 119–130.
- Frankland, P. W., Cestari, V., Filipkowski, R. K., McDonald, R. J., and Silva, A. J. (1998). The dorsal hippocampus is essential for context discrimination but not for contextual conditioning. *Behav. Neurosci.* 112, 863–874.
- Frohardt, R. J., Guarraci, F. A., and Bouton, M. E. (2000). The effects of neurotoxic hippocampal lesions on two effects of context after fear extinction. *Behav. Neurosci.* 114, 227–240.
- Goosens, K. A., and Maren, S. (2001). Contextual and auditory fear conditioning are mediated by the lateral, basal, and central amygdaloid nuclei in rats. *Learn. Mem.* 8, 148–155.
- Hall, G., and Honey, R. C. (1990). Context-specific conditioning in the conditioned-emotional-response procedure. *J. Exp. Psychol. Anim. Behav. Process.* 16, 271–278.
- Hammer, G. D., and Kapp, B. S. (1986). The effects of naloxone administered into the periaqueductal gray on shock-elicited freezing behavior in the rat. *Behav. Neural Biol.* 46, 189–195.
- Han, J. H., Kushner, S. A., Yiu, A. P., Cole, C. J., Matynia, A., Brown, R. A., Neve, R. L., Guzowski, J. F., Silva, A. J., and Josselyn, S. A. (2007). Neuronal competition and selection during memory formation. *Science* 316, 457–460.
- Harris, J. A., Jones, M. L., Bailey, G. K., and Westbrook, R. F. (2000). Contextual control over conditioned responding in an extinction paradigm. *J. Exp. Psychol. Anim. Behav. Process.* 26, 174–185.
- Haubensak, W., Kunwar, P. S., Cai, H., Ciochi, S., Wall, N. R., Ponnusamy, R., Biag, J., Dong, H. W., Deisseroth, K., Callaway, E. M., Fanselow, M. S., Luthi, A., and Anderson, D. J. (2010). Genetic dissection of an amygdala microcircuit that gates conditioned fear. *Nature* 468, 270–276.
- Helmstetter, F. J. (1992). Contribution of the amygdala to learning and performance of conditional fear. *Physiol. Behav.* 51, 1271–1276.
- Helmstetter, F. J., and Fanselow, M. S. (1989). Differential second-order aversive conditioning using contextual stimuli. *Anim. Learn. Behav.* 17, 205–212.
- Helmstetter, F. J., and Landeira-Fernandez, J. (1990). Conditional hypoalgesia is attenuated by naltrexone applied to the periaqueductal gray. *Brain Res.* 537, 88–92.
- Herrero, M. T., Insausti, R., and Gonzalo, L. M. (1991). Cortically projecting cells in the periaqueductal gray matter of the rat. A retrograde fluorescent tracer study. *Brain Res.* 543, 201–212.
- Herry, C., Ciochi, S., Senn, V., Demmou, L., Muller, C., and Luthi, A. (2008). Switching on and off fear by distinct neuronal circuits. *Nature* 454, 600–606.
- Herry, C., Trifilieff, P., Micheau, J., Luthi, A., and Mons, N. (2006). Extinction of auditory fear conditioning requires MAPK/ERK activation in the basolateral amygdala. *Eur. J. Neurosci.* 24, 261–269.
- Hestrin, S., and Galarreta, M. (2005). Electrical synapses define networks of neocortical GABAergic neurons. *Trends Neurosci.* 28, 304–309.
- Heth, C. D., and Rescorla, R. A. (1973). Simultaneous and backward fear conditioning in the rat. *J. Comp. Physiol. Psychol.* 82, 434–443.
- Hitchcock, J., and Davis, M. (1986). Lesions of the amygdala, but not of the cerebellum or red nucleus, block conditioned fear as measured with the potentiated startle paradigm. *Behav. Neurosci.* 100, 11–22.
- Huang, Y. Y., and Kandel, E. R. (1998). Postsynaptic induction and PKA-dependent expression of LTP in the lateral amygdala. *Neuron* 21, 169–178.
- Iordanova, M. D. (2010). Dopamine transmission in the amygdala modulates surprise in an aversive blocking paradigm. *Behav. Neurosci.* 124, 780–788.
- Iordanova, M. D., McNally, G. P., and Westbrook, R. F. (2006). Opioid receptors in the nucleus accumbens regulate attentional learning in the blocking paradigm. *J. Neurosci.* 26, 4036–4045.
- Ji, D., and Wilson, M. A. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nat. Neurosci.* 10, 100–107.
- Ji, J., and Maren, S. (2005). Electrolytic lesions of the dorsal hippocampus disrupt renewal of conditional fear after extinction. *Learn. Mem.* 12, 270–276.
- Jimenez, S. A., and Maren, S. (2009). Nuclear disconnection within the amygdala reveals a direct pathway to fear. *Learn. Mem.* 16, 766–768.
- Johansen, J. P., Hamanaka, H., Diaz-Mataix, L., and Ledoux, J. E. (2010a). Hebbian and neuromodulatory mechanisms act synergistically to instruct associative memory formation. *Soc. Neurosci. Abstr.* (2010 Neuroscience Meeting Planner Online) 914.915.
- Johansen, J. P., Tarpley, J. W., Ledoux, J. E., and Blair, H. T. (2010b). Neural substrates for expectation-modulated fear learning in the amygdala and periaqueductal gray. *Nat. Neurosci.* 13, 979–986.
- Johnston, D., Hoffman, D. A., Colbert, C. M., and Magee, J. C. (1999). Regulation of back-propagating action potentials in hippocampal neurons. *Curr. Opin. Neurobiol.* 9, 288–292.
- Jorntell, H., and Hansel, C. (2006). Synaptic memories upside down: bidirectional plasticity at cerebellar parallel fiber-Purkinje cell synapses. *Neuron* 52, 227–238.
- Kamin, L. J. (1969). “Predictability, surprise, attention, and conditioning,” in *Punishment: A Symposium*, ed. B. C. R. Church (New York: Appleton-Century-Crofts), 276–296.
- Kantak, S. S., Sullivan, K. J., Fisher, B. E., Knowlton, B. J., and Winstein, C. J. (2010). Neural substrates of motor memory consolidation depend on practice structure. *Nat. Neurosci.* 13, 923–925.
- Kiernan, M., and Cranney, J. (1992). Immediate-startle stimulus presentation fails to condition freezing responses to contextual cues. *Behav. Neurosci.* 106, 121–124.
- Kim, J. J., and Fanselow, M. S. (1992). Modality-specific retrograde amnesia of fear. *Science* 256, 675–677.
- Kim, J. J., Krupa, D. J., and Thompson, R. F. (1998). Inhibitory cerebello-olivary projections and blocking effect in classical conditioning. *Science* 279, 570–573.
- Landeira-Fernandez, J., Decola, J. P., Kim, J. J., and Fanselow, M. S. (2006). Immediate shock deficit in fear conditioning: effects of shock manipulations. *Behav. Neurosci.* 120, 873–879.

- Lanuza, E., Nader, K., and Ledoux, J. E. (2004). Unconditioned stimulus pathways to the amygdala: effects of posterior thalamic and cortical lesions on fear conditioning. *Neuroscience* 125, 305–315.
- Lebron, K., Milad, M. R., and Quirk, G. J. (2004). Delayed recall of fear extinction in rats with lesions of ventral medial prefrontal cortex. *Learn. Mem.* 11, 544–548.
- LeDoux, J. E., Iwata, J., Cicchetti, P., and Reis, D. J. (1988). Different projections of the central amygdaloid nucleus mediate autonomic and behavioral correlates of conditioned fear. *J. Neurosci.* 8, 2517–2529.
- Leising, K. J., Sawa, K., and Blaisdell, A. P. (2007). Temporal integration in Pavlovian appetitive conditioning in rats. *Learn. Behav.* 35, 11–18.
- Lev-Ram, V., Mehta, S. B., Kleinfeld, D., and Tsien, R. Y. (2003). Reversing cerebellar long-term depression. *Proc. Natl. Acad. Sci. U.S.A.* 100, 15989–15993.
- Leyland, C. M., and Mackintosh, N. J. (1978). Blocking of first- and second-order autoshaping in pigeons. *Anim. Learn. Behav.* 6, 391–394.
- Li, G., Nair, S. S., and Quirk, G. J. (2009). A biologically realistic network model of acquisition and extinction of conditioned fear associations in lateral amygdala neurons. *J. Neurophysiol.* 101, 1629–1646.
- Lin, C. H., Yeh, S. H., Lu, H. Y., and Gean, P. W. (2003). The similarities and diversities of signal pathways leading to consolidation of conditioning and consolidation of extinction of fear memory. *J. Neurosci.* 23, 8310–8317.
- Lindquist, D. H., Jarrard, L. E., and Brown, T. H. (2004). Perirhinal cortex supports delay fear conditioning to rat ultrasonic social signals. *J. Neurosci.* 24, 3610–3617.
- Mackintosh, N. J. (1975). A theory of attention: variations in the associability of stimuli with reinforcement. *Psychol. Rev.* 82, 276–298.
- Magleby, K. L. (2004). Modal gating of NMDA receptors. *Trends Neurosci.* 27, 231–233.
- Mahanty, N. K., and Sah, P. (1998). Calcium-permeable AMPA receptors mediate long-term potentiation in interneurons in the amygdala. *Nature* 394, 683–687.
- Mahoney, W., and Ayres, J. J. B. (1976). One-trial simultaneous and backward fear conditioning as reflected in conditioned suppression of licking in rats. *Anim. Learn. Behav.* 4, 357–362.
- Maren, S., Aharonov, G., and Fanselow, M. S. (1997). Neurotoxic lesions of the dorsal hippocampus and Pavlovian fear conditioning in rats. *Behav. Brain Res.* 88, 261–274.
- Marlin, N. A. (1983). Second-order conditioning using a contextual stimulus as S1. *Anim. Learn. Behav.* 11, 290–294.
- Marr, D. (1971). Simple memory: a theory for archicortex. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 262, 23–81.
- Marsicano, G., Wotjak, C. T., Azad, S. C., Bisogno, T., Rammes, G., Cascio, M. G., Hermann, H., Tang, J., Hofmann, C., Zieglansberger, W., Di Marzo, V., and Lutz, B. (2002). The endogenous cannabinoid system controls extinction of aversive memories. *Nature* 418, 530–534.
- Matus-Amat, P., Higgins, E. A., Barrientos, R. M., and Rudy, J. W. (2004). The role of the dorsal hippocampus in the acquisition and retrieval of context memory representations. *J. Neurosci.* 24, 2431–2439.
- Matzel, L. D., Held, F. P., and Miller, R. M. (1988). Information and expression of simultaneous and backward associations: implications for contiguity theory. *Learn. Motiv.* 19, 317–344.
- McClelland, J. L., McNaughton, B. L., and O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from connectionist models of learning and memory. *Psychol. Rev.* 102, 419–457.
- McDonald, A. J. (1998). Cortical pathways to the mammalian amygdala. *Prog. Neurobiol.* 55, 257–332.
- McNally, G. P. (2005). Facilitation of fear extinction by midbrain periaqueductal gray infusions of RB101(S), an inhibitor of enkephalin-degrading enzymes. *Behav. Neurosci.* 119, 1672–1677.
- McNally, G. P., Pigg, M., and Weidemann, G. (2004a). Blocking, unblocking, and overexpectation of fear: a role for opioid receptors in the regulation of Pavlovian association formation. *Behav. Neurosci.* 118, 111–120.
- McNally, G. P., Pigg, M., and Weidemann, G. (2004b). Opioid receptors in the midbrain periaqueductal gray regulate extinction of Pavlovian fear conditioning. *J. Neurosci.* 24, 6912–6919.
- McNally, G. P., and Westbrook, R. F. (2006). Predicting danger: the nature, consequences, and neural mechanisms of predictive fear learning. *Learn. Mem.* 13, 245–253.
- Medina, J. F., Nores, W. L., and Mauk, M. D. (2002). Inhibition of climbing fibres is a signal for the extinction of conditioned eyelid responses. *Nature* 416, 330–333.
- Milad, M. R., and Quirk, G. J. (2002). Neurons in medial prefrontal cortex signal memory for fear extinction. *Nature* 420, 70–74.
- Miller, E. K., and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202.
- Miyashita, Y., and Chang, H. S. (1988). Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature* 331, 68–70.
- Moita, M. A., Rosis, S., Zhou, Y., Ledoux, J. E., and Blair, H. T. (2003). Hippocampal place cells acquire location-specific responses to the conditioned stimulus during auditory fear conditioning. *Neuron* 37, 485–497.
- Morgan, M. A., Schulkin, J., and Ledoux, J. E. (2003). Ventral medial prefrontal cortex and emotional perseveration: the memory for prior extinction training. *Behav. Brain Res.* 146, 121–130.
- Mowrer, O. H., and Aiken, E. G. (1954). Contiguity vs. drive-reduction in conditioned fear: temporal variations in conditioned and unconditioned stimulus. *Am. J. Psychol.* 67, 26–38.
- Nader, K., Majidishad, P., Amorapanth, P., and Ledoux, J. E. (2001). Damage to the lateral and central, but not other, amygdaloid nuclei prevents the acquisition of auditory fear conditioning. *Learn. Mem.* 8, 156–163.
- Nieuwenhuis, I. L., and Takashima, A. (2011). The role of the ventromedial prefrontal cortex in memory consolidation. *Behav. Brain Res.* 218, 325–334.
- Onishi, B. K., and Xavier, G. F. (2010). Contextual, but not auditory, fear conditioning is disrupted by neurotoxic selective lesion of the basal nucleus of amygdala in rats. *Neurobiol. Learn. Mem.* 93, 165–174.
- O'Reilly, R. C., and McClelland, J. L. (1994). Hippocampal conjunctive encoding, storage, and recall: avoiding a trade-off. *Hippocampus* 4, 661–682.
- Pare, D., Quirk, G. J., and Ledoux, J. E. (2004). New vistas on amygdala networks in conditioned fear. *J. Neurophysiol.* 92, 1–9.
- Pare, D., Royer, S., Smith, Y., and Lang, E. J. (2003). Contextual inhibitory gating of impulse traffic in the intra-amygdaloid network. *Ann. N. Y. Acad. Sci.* 985, 78–91.
- Pare, D., Smith, Y., and Pare, J. F. (1995). Intra-amygdaloid projections of the basolateral and basomedial nuclei in the cat: *Phaseolus vulgaris*-leucoagglutinin anterograde tracing at the light and electron microscopic level. *Neuroscience* 69, 567–583.
- Parsons, R. G., Gafford, G. M., and Helmstetter, F. J. (2010). Regulation of extinction-related plasticity by opioid receptors in the ventrolateral periaqueductal gray matter. *Front. Behav. Neurosci.* 4, 1–11.
- Parsons, T. C., and Otto, T. (2008). Temporary inactivation of dorsal hippocampus attenuates explicitly nonspatial, unimodal, contextual fear conditioning. *Neurobiol. Learn. Mem.* 90, 261–268.
- Pearce, J. M., and Hall, G. (1980). A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol. Rev.* 87, 532–552.
- Perrett, D. I., Hietanen, J. K., Oram, M. W., and Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 335, 23–30.
- Pezze, M. A., and Feldon, J. (2004). Mesolimbic dopaminergic pathways in fear conditioning. *Prog. Neurobiol.* 74, 301–320.
- Phillips, R. G., and LeDoux, J. E. (1992). Differential contribution of amygdala and hippocampus to cued and contextual fear conditioning. *Behav. Neurosci.* 106, 274–285.
- Pitkanen, A., Pikkarainen, M., Nurminen, N., and Ylinen, A. (2000). Reciprocal connections between the amygdala and the hippocampal formation, perirhinal cortex, and postrhinal cortex in rat. A review. *Ann. N. Y. Acad. Sci.* 911, 369–391.
- Pitkanen, A., Stefanacci, L., Farb, C. R., Go, G. G., Ledoux, J. E., and Amaral, D. G. (1995). Intrinsic connections of the rat amygdaloid complex: projections originating in the lateral nucleus. *J. Comp. Neurol.* 356, 288–310.
- Quinn, J. J., Loya, F., Ma, Q. D., and Fanselow, M. S. (2005). Dorsal hippocampus NMDA receptors differentially mediate trace and contextual fear conditioning. *Hippocampus* 15, 665–674.
- Quinn, J. J., Ma, Q. D., Tinsley, M. R., Koch, C., and Fanselow, M. S. (2008a). Inverse temporal contributions of the dorsal hippocampus and medial prefrontal cortex to the expression of long-term fear memories. *Learn. Mem.* 15, 368–372.
- Quinn, J. J., Wied, H. M., Ma, Q. D., Tinsley, M. R., and Fanselow, M. S. (2008b). Dorsal hippocampus involvement in delay fear conditioning depends upon the strength of the tone-footshock association. *Hippocampus* 18, 640–654.
- Quirk, G. J., Russo, G. K., Barron, J. L., and Lebron, K. (2000). The role of ventromedial prefrontal cortex in the recovery of extinguished fear. *J. Neurosci.* 20, 6225–6231.
- Raineki, C., Holman, P. J., Debiec, J., Bugg, M., Beasley, A., and Sullivan, R. M. (2010). Functional emergence

- of the hippocampus in context fear learning in infant rats. *Hippocampus* 20, 1037–1046.
- Repa, J. C., Muller, J., Apergis, J., Desrochers, T. M., Zhou, Y., and Ledoux, J. E. (2001). Two different lateral amygdala cell populations contribute to the initiation and storage of memory. *Nat. Neurosci.* 4, 724–731.
- Rescorla, R. A. (1980). Simultaneous and successive associations in sensory preconditioning. *J. Exp. Psychol. Anim. Behav. Process.* 6, 207–216.
- Rescorla, R. A. (2002). “Extinction,” in *Psychology at the Turn of the Millennium: Vol. 1. Cognitive, Biological, and Health Perspectives*, eds L. Backman and C. Von Hofsten (Hove: Taylor and Francis), 217–244.
- Rescorla, R. A. (2003). Protection from extinction. *Learn. Behav.* 31, 124–132.
- Richmond, M. A., Yee, B. K., Pouzet, B., Veenman, L., Rawlins, J. N., Feldon, J., and Bannerman, D. M. (1999). Dissociating context and space within the hippocampus: effects of complete, dorsal, and ventral excitotoxic hippocampal lesions on conditioned freezing and spatial learning. *Behav. Neurosci.* 113, 1189–1203.
- Rodrigues, S. M., Schafe, G. E., and Ledoux, J. E. (2001). Intra-amygdala blockade of the NR2B subunit of the NMDA receptor disrupts the acquisition but not the expression of fear conditioning. *J. Neurosci.* 21, 6889–6896.
- Rogan, M. T., and Ledoux, J. E. (1995). LTP is accompanied by commensurate enhancement of auditory-evoked responses in a fear conditioning circuit. *Neuron* 15, 127–136.
- Royer, S., Martina, M., and Pare, D. (1999). An inhibitory interface gates impulse traffic between the input and output stations of the amygdala. *J. Neurosci.* 19, 10575–10583.
- Royer, S., and Pare, D. (2002). Bidirectional synaptic plasticity in intercalated amygdala neurons and the extinction of conditioned fear responses. *Neuroscience* 115, 455–462.
- Rudy, J. W. (2009). Context representations, context functions, and the parahippocampal-hippocampal system. *Learn. Mem.* 16, 573–585.
- Rudy, J. W., Barrientos, R. M., and O’Reilly, R. C. (2002). Hippocampal formation supports conditioning to memory of a context. *Behav. Neurosci.* 116, 530–538.
- Rudy, J. W., Biedenkapp, J. C., and O’Reilly, R. C. (2005). Prefrontal cortex and the organization of recent and remote memories: an alternative view. *Learn. Mem.* 12, 445–446.
- Rudy, J. W., and O’Reilly, R. C. (2001). Conjunctive representations, the hippocampus, and contextual fear conditioning. *Cogn. Affect. Behav. Neurosci.* 1, 66–82.
- Rumpel, S., Ledoux, J., Zador, A., and Malinow, R. (2005). Postsynaptic receptor trafficking underlying a form of associative learning. *Science* 308, 83–88.
- Sacchetti, B., Lorenzini, C. A., Baldi, E., Tassoni, G., and Bucherelli, C. (1999). Auditory thalamus, dorsal hippocampus, basolateral amygdala, and perirhinal cortex role in the consolidation of conditioned freezing to context and to acoustic conditioned stimulus in the rat. *J. Neurosci.* 19, 9570–9578.
- Santini, E., Ge, H., Ren, K., Pena De Ortiz, S., and Quirk, G. J. (2004). Consolidation of fear extinction requires protein synthesis in the medial prefrontal cortex. *J. Neurosci.* 24, 5704–5710.
- Santini, E., Muller, R. U., and Quirk, G. J. (2001). Consolidation of extinction learning involves transfer from NMDA-independent to NMDA-dependent memory. *J. Neurosci.* 21, 9009–9017.
- Savander, V., Go, C. G., Ledoux, J. E., and Pitkanen, A. (1995). Intrinsic connections of the rat amygdaloid complex: projections originating in the basal nucleus. *J. Comp. Neurol.* 361, 345–368.
- Schenberg, E. E., and Oliveira, M. G. (2008). Effects of pre or posttraining dorsal hippocampus D-AP5 injection on fear conditioning to tone, background, and foreground context. *Hippocampus* 18, 1089–1093.
- Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.
- Schultz, W., and Dickinson, A. (2000). Neuronal coding of prediction errors. *Annu. Rev. Neurosci.* 23, 473–500.
- Selden, N. R., Everitt, B. J., Jarrard, L. E., and Robbins, T. W. (1991). Complementary roles for the amygdala and hippocampus in aversive conditioning to explicit and contextual cues. *Neuroscience* 42, 335–350.
- Shi, C., and Davis, M. (1999). Pain pathways involved in fear conditioning measured with fear-potentiated startle: lesion studies. *J. Neurosci.* 19, 420–430.
- Shipley, R. H. (1974). Extinction of conditioned fear in rats as a function of several parameters of CS exposure. *J. Comp. Physiol. Psychol.* 87, 699–707.
- Sierra-Mercado, D. Jr., Corcoran, K. A., Lebron-Milad, K., and Quirk, G. J. (2006). Inactivation of the ventromedial prefrontal cortex reduces expression of conditioned fear and impairs subsequent recall of extinction. *Eur. J. Neurosci.* 24, 1751–1758.
- Skaggs, W. E., and McNaughton, B. L. (1992). Computational approaches to hippocampal function. *Curr. Opin. Neurobiol.* 2, 209–211.
- Sotres-Bayon, F., Bush, D. E., and Ledoux, J. E. (2007). Acquisition of fear extinction requires activation of NR2B-containing NMDA receptors in the lateral amygdala. *Neuropsychopharmacology* 32, 1929–1940.
- Stiedl, O., Birkenfeld, K., Palve, M., and Spiess, J. (2000). Impairment of conditioned contextual fear of C57BL/6J mice by intracerebral injections of the NMDA receptor antagonist APV. *Behav. Brain Res.* 116, 157–168.
- Stote, D. L., and Fanselow, M. S. (2004). NMDA receptor modulation of incidental learning in Pavlovian context conditioning. *Behav. Neurosci.* 118, 253–257.
- Sutton, R. S., and Barto, A. G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. *Psychol. Rev.* 88, 135–170.
- Sutton, R. S., and Barto, A. G. (1990). “Time-derivative models of Pavlovian reinforcement,” in *Learning and Computational Neuroscience: Foundations of Adaptive Networks*, eds M. Gabriel and J. Moore (Cambridge, MA: MIT Press), 497–537.
- Suzuki, W. A. (2010). Untangling memory from perception in the medial temporal lobe. *Trends Cogn. Sci. (Regul. Ed.)* 14, 195–200.
- Thomas, B. L., Larsen, L., and Ayres, J. B. (2003). Role of context similarity in ABA, ABC, and AAB renewal paradigms: implications for theories of renewal and for treating human phobias. *Learn. Motiv.* 34, 410–436.
- Tomaz, C., Dickinson-Anson, H., Mcgaugh, J. L., Souza-Silva, M. A., Viana, M. B., and Graeff, F. G. (1993). Localization in the amygdala of the amnesic action of diazepam on emotional memory. *Behav. Brain Res.* 58, 99–105.
- Tully, K., Li, Y., Tsvetkov, E., and Bolshakov, V. Y. (2007). Norepinephrine enables the induction of associative long-term potentiation at thalamo-amygdala synapses. *Proc. Natl. Acad. Sci. U.S.A.* 104, 14146–14150.
- Vlachos, I., Herry, C., Luthi, A., Aersen, A., and Kumar, A. (2011). Context-dependent encoding of fear and extinction memories in a large-scale network model of the basal amygdala. *PLoS Comput. Biol.* 7, e1001104. doi: 10.1371/journal.pcbi.1001104
- Vu, E. T., and Krasne, F. B. (1992). Evidence for a computational distinction between proximal and distal neuronal inhibition. *Science* 255, 1710–1712.
- Wilensky, A. E., Schafe, G. E., Kristensen, M. P., and Ledoux, J. E. (2006). Rethinking the fear circuit: the central nucleus of the amygdala is required for the acquisition, consolidation, and expression of Pavlovian fear conditioning. *J. Neurosci.* 26, 12387–12396.
- Willis, W. D. J., and Westlund, K. N. (2004). “Pain system,” in *The Human Nervous System*, 2nd Edn, eds G. Paxinos and J. Mai (San Diego, CA: Elsevier), 1125–1170.
- Wilson, A., Brooks, D. C., and Bouton, M. E. (1995). The role of the rat hippocampal system in several effects of context in extinction. *Behav. Neurosci.* 109, 828–836.
- Wilson, M. A., and McNaughton, B. L. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science* 265, 676–679.
- Wiltgen, B. J., Sanders, M. J., Anagnostaras, S. G., Sage, J. R., and Fanselow, M. S. (2006). Context fear learning in the absence of the hippocampus. *J. Neurosci.* 26, 5484–5491.
- Won, J., and Silva, A. J. (2008). Molecular and cellular mechanisms of memory allocation in neuron networks. *Neurobiol. Learn. Mem.* 89, 285–292.
- Woodruff, A. R., and Sah, P. (2007). Networks of parvalbumin-positive interneurons in the basolateral amygdala. *J. Neurosci.* 27, 553–563.
- Yaniv, D., Desmedt, A., Jaffard, R., and Richter-Levin, G. (2004). The amygdala and appraisal processes: stimulus and response complexity as an organizing factor. *Brain Res. Brain Res. Rev.* 44, 179–186.
- Young, S. L., Bohenek, D. L., and Fanselow, M. S. (1994). NMDA processes mediate anterograde amnesia of contextual fear conditioning induced by hippocampal damage: immunization against amnesia by context preexposure. *Behav. Neurosci.* 108, 19–29.
- Zelikowsky, M., Bissiere, S., Abdipranato, A., Hast, T., Vissel, B., and Fanselow, M. S. (2010). Contextual fear formed in the absence of the hippocampus is fleeting and requires infralimbic cortex. *Soc. Neurosci. Abstr.* (2010 Neuroscience Meeting Planner Online) 915.914.
- Zelikowsky, M., Pham, D. L., and Fanselow, M. S. (2011). Temporal factors control hippocampal contributions to fear renewal after extinction. *Hippocampus* doi: 10.1002/hipo.20954. [Epub ahead of print].
- Zhou, Y., Won, J., Karlsson, M. G., Zhou, M., Rogerson, T., Balaji, J., Neve, R., Poirazi, P., and Silva, A. J. (2009). CREB

regulates excitability and the allocation of memory to subsets of neurons in the amygdala. *Nat. Neurosci.* 12, 1438–1443.

Zimmerman, J. M., and Maren, S. (2010). NMDA receptor antagonism in the basolateral but not central amygdala blocks the extinction of Pavlovian fear

conditioning in rats. *Eur. J. Neurosci.* 31, 1664–1670.

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

Received: 30 March 2011; accepted: 07 July 2011; published online: 26 July 2011.

Citation: Krasne FB, Fanselow MS and Zelikowsky M (2011) Design of a neurally plausible model of fear learning. *Front. Behav. Neurosci.* 5:41. doi: 10.3389/fnbeh.2011.00041

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