Published in final edited form as:

Trends Neurosci. 2011 June ; 34(6): 283–292. doi:10.1016/j.tins.2011.03.005.

Placing prediction into the fear circuit

Gavan P. McNally^{1,*}, Joshua P. Johansen², and Hugh T. Blair³

¹ School of Psychology, The University of New South Wales, Sydney, NSW, Australia.

² Center for Neural Science, New York University, New York, New York, U.S.A.

³ Department of Psychology, University of California Los Angeles, Los Angeles, California, USA

Abstract

Pavlovian fear conditioning depends on synaptic plasticity at amygdala neurons. Here we review recent electrophysiological, molecular, and behavioral evidence suggesting the existence of a distributed neural circuitry regulating amygdala synaptic plasticity during fear learning. This circuitry, which involves projections from the midbrain periaqueductal gray (PAG) region, can be linked to prediction error and expectation modulation of fear learning as described by associative and computational learning models. It controls whether, and how much, fear learning occurs by signalling aversive events when they are unexpected. Functional neuroimaging and clinical studies indicate that this prediction circuit is recruited in humans during fear learning and contributes to exposure-based treatments for clinical anxiety. This aversive prediction error circuit may represent a conserved mechanism for regulating fear learning in mammals.

Introduction

Pavlovian fear conditioning involves pairing of a conditioned stimulus (CS) with an aversive unconditioned stimulus (US), such as a footshock. As a result of these pairings, subjects express a diverse but co-ordinated range of conditioned responses (eg. changes in heart rate, respiration, blood pressure, and species-specific defense responses) to the CS on subsequent presentations [1, 2]. Significant progress has been made in understanding the neural mechanisms for this learning. Acquisition of fear learning depends on the lateral amygdala (LA) whereas expression of conditioned fear depends on the central amygdala (CeA) and its projections to the midbrain, brainstem and hypothalamic nuclei [3 - 6] (**Box 1**). This circuitry for fear learning and memory formation is well preserved across a variety of species and has become a primary focus of research into the neurobiology of human anxiety disorders [7, 8]. Here we review recent findings suggesting that neural plasticity in the amygdala may be supervised by neural circuitry originating from the midbrain PAG. Such a

Crown Copyright © 2011 Published by Elsevier Ltd. All rights reserved.

* Corresponding author: G.P. McNally (g.mcnally@unsw.edu.au).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

pathway is instrumental in generating an instructive “teaching” signal that contributes to the modulation of synaptic plasticity during fear conditioning.

Modulation of learning by expectation: Prediction errors as teaching signals

Pavlovian fear conditioning depends on the potentiation of CS input synapses onto LA neurons [3, 4]. Such plasticity is triggered by afferent pathways that transmit US related information to LA neurons. Many different CSs can come to elicit fear responses after being paired with an aversive US, so it is natural to regard these pathways as carrying a teaching signal which instructs learning, and synaptic plasticity, across CS – US pairings. Aversive USs may act as teaching signals to trigger plasticity at CS input synapses to the LA, at least in part, by causing depolarization and action potential firing in LA neurons while CS inputs are active [9, 10]. There is reason to believe that the strength of this teaching signal is not invariant; rather, it is modulated by the expectation of the US during each learning trial. Several lines of evidence show that Pavlovian fear conditioning is more effective when the CS is paired with an unexpected US than with an expected US [11 – 14]. For example, the acquisition of fear to a CS is negatively accelerated across learning trials, so that fear of a CS increases most during early CS – US pairings (when the US is unexpected) and least during later pairings (when the CS has come to predict the US). To explain such findings, learning theories have posited that fear conditioning is not instructed by a simple sensory representation of the US, but instead by an error signal measuring the difference between the US actually present and that expected. In the following sections, we briefly review three types of error signals that have been proposed by formal learning theories.

The Rescorla-Wagner learning rule—The Rescorla-Wagner learning rule [11] proposes that learning is controlled by an error signal that encodes the difference between the actual versus expected intensity of the US. This error signal dictates variations in the effectiveness of the US in supporting learning. If the actual US is denoted as λ and the expected US as ΣV (to indicate the summed associative strengths, V , of all CSs preceding the US), then the error signal is computed as $\lambda - \Sigma V$. The learning rule for synaptic modification and change (ΔV) in associative strength under these conditions is:

$$\Delta V = S(\lambda - \Sigma V)$$

where S is a learning rate parameter. If a US occurs unexpectedly, then the actual US will exceed that expected ($\lambda > \Sigma V$), and a positive prediction error generated to drive synaptic plasticity and fear learning. By contrast, if the occurrence of the US is expected and matches expectations ($\lambda = \Sigma V$), then the error signal is zero and no synaptic plasticity or fear learning occurs. If the actual US is less than expected ($\lambda < \Sigma V$), then the error signal is negative and generates an instructive signal for extinction learning which reduces fear of the CS.

The Pearce-Hall learning rule—The Pearce-Hall learning rule [15] posits that an error signal regulates the amount of attention paid to the CS on each conditioning trial. A CS commands attention if it is a poor predictor of the US. Specifically, attention (α) to the CS on the current trial (n) is proportional to the prediction error on the previous trial ($n - 1$). Thus,

$$\alpha_n = |\lambda - \Sigma V|_{n-1},$$

and the instructive signal for modifying synaptic plasticity and associative strength is:

$$\Delta V = \alpha_n S - \lambda.$$

If the CS was a poor predictor of the US on the previous trial, then α is large on the following trial and the instructive signal will be high. By contrast, if the CS was a good predictor of the US during the previous trial, then α will be small on the following trial and the signal will be too small to strengthen the association. In this way, learning occurs preferentially to CSs whose consequences are uncertain.

The temporal-difference learning rule—The temporal-difference learning rule [12] does not incorporate an error signal that computes the difference between actual versus expected US intensities. Instead, the TD error signal sums the actual and expected US intensities together, and then compares the momentary value of this sum (which may be denoted as $\lambda_t + \Sigma V_t$ at time t) against the prior moment's value of the expected US intensity (denoted here as ΣV_{t-1}). Thus, the TD error signal may be written:

$$\Delta V = S (\lambda_t + \Sigma V_t - \Sigma V_{t-1}).$$

Note that the TD error signal arises from a comparison that is made across successive moments in time, t versus $t-1$ (hence the name, temporal difference learning). The essence of the TD rule is that learning is directly driven by moments of surprise, which are defined as moments when either the actual or expected US intensity (or the sum of both) exceed the US intensity than what was expected just a moment ago.

The Rescorla-Wagner and TD learning rules (but not the Pearce-Hall rule) rely upon signed prediction error signals, which can be either positive or negative depending upon circumstances. To encode these signed prediction errors, neurons could increase their firing rates when error is positive, and decrease their firing rates when error is negative. Neurobiological evidence indicates that prediction error signals may instruct several well-studied forms of learning, including cerebellar motor learning [16], developmental plasticity in the avian inferior colliculus [17], and reward learning mediated by the midbrain dopamine (DA) system [18, 19] (**Box 2**). Until recently little was known regarding the neural representations of prediction error signals in fear learning.

Instruction of fear acquisition by aversive prediction errors

During fear conditioning, memories for the CS - US association are thought to be stored by synaptic plasticity in LA neurons, and studies have shown that LA neurons respond preferentially to an unexpected rather than expected US [20, 21]. This suggests that LA neurons may receive instructive teaching inputs that encode an aversive prediction error signal, but if so, from where might this teaching signal derive?

Several studies suggest that instructive prediction error signals may arise from the midbrain periaqueductal gray (PAG), a structure that has been implicated in the expression of fear behavior (**Box 3**) as well as the regulation of aversive stimulus processing. Direct stimulation of PAG neurons can serve as a US in the absence of a peripheral shock to support fear conditioning [22], supporting the view that the PAG is positioned to play a role in instructing associative plasticity during fear conditioning. In a recent study [21] (**Figure 1**), neurons were recorded from LA and PAG neurons during Pavlovian fear conditioning in rats, and the PAG was inactivated while recording from LA neurons. Pharmacological inhibition of PAG neurons during fear acquisition prevented learning. Critically, shock US-evoked responding in LA and PAG neurons were modulated by expectation. Across the course of auditory CS – shock US pairings, US-evoked responses in LA and PAG neurons declined, concomitant with an increase in expression of conditioned fear responses. This suggests that US-evoked responses of LA and PAG neurons declined as the shock became expected and prediction error declined. This interpretation was supported by the finding that in well-trained rats, neurons recorded in the LA and PAG responded more to the US when it was presented by itself (that is, unexpectedly) than when it was signalled by the CS, and thus expected to occur. Finally, there was evidence that the expectancy modulated US-evoked response in LA neurons is relayed from the PAG to the LA, because PAG inactivation reduced US-evoked responding in LA neurons. US-evoked depolarization and action potential firing in LA neurons are likely to be critical components of the instructive signal that triggers plasticity at CS input synapses during fear conditioning (9, 10). Thus these data suggest a plausible neural mechanism for limiting learning when the CS predicts the US.

Endogenous opioids acting at μ -opioid receptors (MOR) in the ventrolateral PAG (vlPAG) are candidate receptors for mediating this mechanism. MOR antagonism augments acquisition of fear learning by removing limits on the ability of an expected US to condition fear [23 - 26]. A similar augmentation of fear conditioning has been observed in human subjects. Specifically, functional magnetic imaging resonance (fMRI) recordings in subjects who underwent fear conditioning revealed that opioid receptor antagonism during the experiment prevented the diminution of amygdala blood oxygen level dependent (BOLD) responses typically observed across repeated CS – US pairings [27]. Studies using more complex behavioral training paradigms in rats have shown that activation of MORs in the vlPAG reduces the effectiveness of an expected US as a reinforcer during fear learning [28, 29]. For example, associative blocking tasks have been used to study the role of vlPAG MOR in learning about expected versus unexpected events. In one such study [28] (**Figure 1**), rats were first trained to fear a CS “A” [ie. CS(A)], via pairings with a shock US. In the second stage of the experiment (Stage II), rats were trained to fear two compounds. One compound consisted of CS(A) + CS(B) paired with the shock US. The second consisted of CS(C)+CS(D) paired with the shock US. The prediction error during the CS(AB)-shock pairings was low, since CS(A) had been previously paired with the shock US in Stage I, and hence, the shock was expected in Stage II. By contrast, the prediction error during CS(CD)-shock pairings was high because neither CS(C) nor CS(D) had previously been paired with the shock US in Stage I, and hence the shock was unexpected during Stage II. Rats were later tested for fear reactions to CS(B) and CS(D). Control rats showed evidence for

associative blocking so that fear was less to CS(B) than to CS(D). That is, the prior fear learning about CS(A) blocked later fear learning to CS(B). This pattern of learning was prevented by antagonising MORs in the vIPAG during Stage II of the experiment. Taken together, these data suggest that the PAG may be part of the teaching signal pathway for fear learning, instructing LA associative plasticity. Furthermore, this function depends, at least in part, on vIPAG MORs.

It is worth considering whether these data also permit selection between the different error signals proposed by formal learning theories. Electrophysiological data show that the US-evoked population response in PAG is largest on the first CS-US pairing trial [21]. This result does not easily permit selection between different error signals because most theories predict that such signals decline across CS – US pairings. Behavioral data show that vIPAG MOR contributions to associative blocking are observed with a single Stage II conditioning trial [29]. This finding is more consistent with error signals described by Rescorla-Wagner and TD models (causing variations in US processing) than the error signals described by the Pearce-Hall model (causing variations in CS processing). Nonetheless, some amygdala neurons encode Pearce-Hall type attentional signals [20, 30, 31] and there is behavioural evidence that amygdala NMDA receptors regulate an attentional or salience signal during fear learning [32]. Significant theoretical effort is being devoted to the development of hybrid associative models, which would allow for both US and CS error signals in Pavlovian learning [33, 34]. It will be of interest to determine whether this effort permits more parsimonious explanation of neuronal activity during fear conditioning.

Instruction of fear extinction by negative prediction errors

For aversive conditioning, a negative prediction error is defined as a signal which is generated when the actual US (λ) is less than expected (V). This error can be generated by increasing expectation (V) (e.g., overexpectation) or by decreasing US intensity (λ). The simplest example is fear extinction, when a CS which was previously paired with an aversive US is subsequently presented alone in the absence of the US. This negative prediction error instructs loss of fear during extinction training.

Fear extinction depends on the LA and the prefrontal cortex (PFC), where activation of NMDA receptors, their associated intracellular signalling cascades, and synaptic plasticity is crucial to extinction learning and memory storage [3, 35 - 37]. If opioid signalling in the PAG contributes to negative feedback regulation during fear learning (as indicated above), then behavioral fear extinction and the plasticity upon which it depends might also be influenced by PAG opioids. Recent studies support this. Systemic [38, 39] or vIPAG microinjections [40 – 42] of MOR antagonists prevent fear extinction learning. Conversely, fear extinction learning can be facilitated by infusions of a peptidase inhibitor that reduces catabolism of vIPAG enkephalins [43]. Moreover, vIPAG infusions of MOR antagonists not only impair extinction learning, they also prevent the normal increase in phosphorylation of the extracellular related kinase/mitogen activated protein kinase (ERK/MAPK) observed in the PFC and amygdala during extinction learning [42] and which has been shown to be critical for fear extinction memory consolidation [44 - 46]. Thus, the vIPAG regulates synaptic plasticity in the LA and PFC during fear extinction learning.

Another line of evidence that opioid receptors, although not necessarily those in the midbrain, are important for learning not to fear is derived from clinical studies. Exposure therapies for human anxieties are modelled on fear extinction training from animal conditioning studies. Just as opioid receptors are critical for fear reduction by extinction training in animal conditioning studies, so too are they important for the therapeutic benefit of exposure therapies for human clinical anxiety. Administrations of opioid antagonists prior to exposure based treatments for simple phobias reduce the efficacy of these treatments [47 – 49]. Moreover, exogenous opiates administered in the hours to days following a traumatic event can reduce development of post-traumatic stress disorder (PTSD) [50].

The Rescorla-Wagner and TD learning rules posit that learning is instructed by a signed error signal. If fear extinction is instructed by this error signal, then neurons encoding prediction errors might be expected to decrease their firing rates during omission of an expected aversive US. Electrophysiological data do not show evidence for phasic responses—inhibitory or excitatory—in firing of PAG or LA neurons upon omission of an expected aversive US [20, 21]. It is possible that negative prediction errors for fear learning are encoded by other neurons. Likewise, it is possible that PAG neurons signal an unsigned error as described by Pearce-Hall, although behavioural data using single trial blocking studies are inconsistent with this. Another possible explanation is simply that the negative component of the fear error signal is difficult to detect. This negative error signal may be smaller in magnitude and more distributed across time than the phasic negative error signals observed in the reward learning system. The signed TD prediction error is the time derivative of expected future reinforcement [12], and in the case of fear learning, this would be the time derivative of fear itself (since fear may be regarded as the expectation of an aversive US). Fear, unlike anticipated reward, is rarely fleeting. Hence, if amygdala plasticity is instructed by a TD-like error signal encoding the time derivative of fear, then rapid phasic responses to the omission of an expected aversive stimulus might not be observed. Instead, smaller and slower signals might instruct the gradual decline of fear during the transition from danger to safety.

Circuit level mechanisms for predicting danger

Computation of aversive prediction error, and expectancy-modulation of US-teaching signals, can be achieved via a negative feedback circuit [51 - 53] (**Figure 2**). This requires convergence between efferents of the CS system and afferents of the US system. Extrapolating to fear conditioning [54, 55], output of the conditioned fear system corresponds to the expected outcome of the conditioning trial because it carries information about the outcome of previous CS – US pairings ($-\Sigma V$), whereas transmission in the somatosensory system conveys information about the actual aversive US on the current trial (λ).

vIPAG is an important locus of neuroanatomical convergence in the pathways conveying information about the actual aversive US (λ) and the expected aversive US (ΣV). vIPAG is a target of ascending nociceptive pathways conveying information about aversive USs present on a conditioning trial [56]. vIPAG also receives extensive projections from PFC and amygdala neurons whose activity determines expression of conditioned fear. Medial central

nucleus (CeAm) projection neurons are excited during expression of conditioned fear [57-59]. The CeAm $\rightarrow$ vIPAG projection could convey information about the expected aversive US ($-\Sigma V$). The PAG also receives projections from the PFC, including prelimbic, infralimbic, and orbital regions [60 – 63], regions implicated in conditioned fear expression [64]. Cortical projections to the PAG play important roles in response selection and coping during stress [62, 65], for example, as a consequence of exposure to inescapable shock [66, 67], and could convey information about the expected aversive US.

The circuit level mechanism(s) by which US processing is inhibited during fear conditioning have not been carefully studied and are largely unknown. CeA neurons are mainly inhibitory and their output is thought to inhibit neurons at efferent target sites [57]. CeA stimulation produces inhibition and excitation of different populations of PAG neurons and these effects are partially blocked by local infusion of a MOR antagonist [68]. CeA-mediated MOR activation may directly dampen or attenuate US processing in vIPAG, so that the ascending output of PAG neurons codes for the error signal ($\lambda - \Sigma V$) [55]. A different possibility is that CeA-mediated recruitment of PAG neurons activates descending antinociceptive circuits which attenuate US processing at the level of the spinal and trigeminal dorsal horn [54, 69, 70]. The relationship between the roles of PAG in encoding a fear teaching signal and fear expression is unclear. The available evidence favours the possibility that these roles are partially distinct. Learning in response to PAG teaching signals can be pharmacologically dissociated from fear expression [28, 29]. Moreover, different populations of PAG neurons have been described [71], including in response to electrical or chemical stimulation of the CeA [68].

Although the PAG likely transmits US teaching signal information to the LA, it does send direct projections there [61]. Thus, an indirect pathway is almost certainly involved. One pathway involves the midline and intralaminar thalamus. Projections from the vIPAG terminate throughout the midline thalamus [72]. These, in turn, project widely to the medial PFC, the orbital frontal cortex (OFC), anterior cingulate (Cg), and rostral agranular insular cortex (RAIC) [72 - 75] and include direct and indirect (from the PFC) [75, 76] projections to amygdala. These projections, part of the ‘medial’ pain pathway [77], are important candidates for conveying US teaching signals to LA. Findings from human neuroimaging as well as rodent functional neuroanatomical studies support this (**Figure 3**). BOLD signals recorded in the human PFC (notably the insula, orbital, and cingulate), are related to the magnitude of prediction error during fear learning [e.g., 78 – 80]. Moreover, the US-evoked BOLD signal in midline thalamus, PFC, and amygdala diminishes across CS – US pairings [27; 82]. There is a negative correlation between US-evoked BOLD signal in these regions and the self-report of US expectancy [82]. In rodents, unexpected, but not expected, CS – aversive US pairings produce activation of midline thalamic regions as well as the PFC and LA [83], including in identified midline thalamus $\rightarrow$ PFC projection neurons [83].

These findings suggest that midline thalamic $\rightarrow$ PFC pathway activity may convey an aversive US teaching signal, however, there is currently little additional behavioural evidence that supports this possibility. Lesion studies have implicated midline thalamus in fear learning. Posterior intralaminar thalamus or parietal insular cortex lesions impair fear acquisition in rats, consistent with a role for these regions in aversive US processing [84],

but conflicting effects have been reported in other studies [85, 86]. Likewise, several studies have implicated PFC in fear learning. Lesions of rodent dorsal PFC, encompassing prelimbic (PL) and cingulate (Cg) cortex, can augment acquisition of conditioned fear [87] as can lesions of the RAIC [88]. Pharmacological activation of glutamate receptors [89] or electrical stimulation [90] as a US in rodent Cg is sufficient to support fear learning, whereas pharmacological inhibition or antagonism of glutamate receptors in the Cg can retard or prevent such learning [89, 91]. These effects could be due to modulation of the US teaching signal. However, other interpretations are possible and conflicting findings have been reported [86, 88, 90 - 92]. A recent demonstration that reversible inactivation of rodent dorsomedial PFC (dmPFC), encompassing caudal Cg and dorsal PL, prevents associative blocking of fear learning [83] and so restores the effectiveness of an expected US as an aversive reinforcer, provides important evidence in support of this. However, further work is needed to understand these circuit level mechanisms. In particular, studies are needed which use behavioral paradigms (e.g., associative blocking [Figure 1]) that permit clear isolation of the contribution of prediction error to fear learning and involve selective, reversible modulation of discrete cortical regions during such learning. Recently developed optogenetic techniques (eg. [10, 93, 94]) could allow a more temporally precise tool with which to parse the contribution of these brain regions to specific events occurring during fear learning.

Conclusions

The ability to use past experience to predict the future, and respond appropriately, is a signature of adaptive behavior. Pavlovian fear conditioning enables learning about, and adaptive responding to, sources of danger. Central to this learning is encoding the predictive relationship between a CS and an aversive US, so that synaptic plasticity and learning occur preferentially to unexpected sources of danger (fear conditioning) and unexpected sources of safety (fear extinction). The evidence reviewed here supports the view that a neural circuit that signals whether aversive events are expected or unexpected imbues amygdala-based learning mechanisms with sensitivity to predictive relations. This circuit involves the midbrain PAG and the actions of the endogenous opioids at MORs therein, and also involves an ascending circuit to the PFC, via the midline thalamus. The functional neuroimaging studies reviewed here indicate that this circuit is recruited in humans during fear learning and contributes to the therapeutic benefit of exposure-based treatments of human clinical anxiety.

Understanding of the fear prediction circuit is nascent, and much remains to be learned (**Box 4**). Nonetheless, the existence of a discrete teaching circuit for fear learning has important implications. Theoretically, identification of the circuits, receptors, and molecules determining variations in US and CS effectiveness during Pavlovian fear conditioning is an important step towards reconciliation of psychological, computational, and neurobiological approaches to the study of fear learning. Practically, understanding these mechanisms may provide new insights into the aetiology, treatment, and prevention of fear and anxiety disorders in humans.

Box 1

Amygdala circuitry involved in fear learning and memory. During Pavlovian fear conditioning, sensory thalamic and cortical afferents carry CS inputs to LA pyramidal neurons (Figure I). Afferents carrying shock US inputs converge onto the same LA neurons. This CS – US convergence initiates synaptic plasticity mediated by postsynaptic NMDA receptors on LA neurons, resulting in a potentiation of CS inputs. The potentiated CS inputs to LA neurons allow expression of fear responses via an intra-amygdala circuitry linking LA to the basal amygdala (BA) and CeA. Outputs from CeA to the hypothalamus, midbrain, and brainstem generate the co-ordinated expression of behavioural and autonomic fear responses.

Box 2. Roles for dopamine in prediction error and fear learning

Midbrain dopamine (DA) neurons code for reward prediction errors. The firing of these neurons conforms to assumptions of associative learning models [18, 19] and their output is thought to serve as a teaching signal instructing plasticity in the striatum. The canonical findings from recordings in primates during Pavlovian appetitive conditioning are that midbrain DA neurons show increases in firing to unexpected rewards, little change in firing to expected rewards, and inhibited firing to omission of an expected reward [18, 19]. Some midbrain DA neurons respond not only to rewards and their signals, but also to aversive USs and their signals. Some primate [99] and rodent [100] DA neurons are inhibited by aversive USs or their CSs, whereas others show phasic excitations. In both species there is neuroanatomical segregation of these two populations. These DA neurons are at least partly sensitive to prediction error because the magnitude of aversive US-elicited phasic excitations and inhibitions decrease as the aversive US becomes expected [99].

The role for DA in prediction errors during fear conditioning depends on the circuits in which its receptors are located. Antagonizing dopamine D1 and D2 receptors in the nucleus accumbens (Acb) prevents associative blocking of fear learning [101]. Similar findings are observed when antagonizing Acb MORs [102]. This role for DA and MORs in the Acb is directly linked to an error signal determining CS associability [101]. Acb DA and MORs regulate the attention paid to a CS as a function of how well that CS predicts its consequences. Thus, a vPAG based circuit may determine variations in US effectiveness whereas a midbrain DA and Acb based circuit may determine variations in CS effectiveness. Nonetheless, this distinction is not absolute because D1 and D2 dopamine receptor activation in the amygdala is involved in fear learning [103 – 105] and antagonizing these receptors prevents associative blocking of fear learning by changing effectiveness of the US as a reinforcer [106].

Lateral habenula neurons also show changes in firing to aversive USs and to CSs which signal them [107]. They show increased firing rates to unexpected aversive USs which decline as the US becomes expected. Habenula neurons likewise show increased firing rates to CSs which predict an aversive US and these same neurons show graded responses to appetitive CSs but with firing rates opposite in sign. Thus, lateral habenula

neurons are most responsive to CSs which signal an aversive US or the absence of an appetitive US and least responsive to CSs which signal the absence of an aversive US or the presence of an appetitive US. This overlap in the neuronal coding of a CS signaling danger and a CS signaling the absence of reward is precisely that anticipated by Konorski [108]. Nonetheless, the role of the lateral habenula in fear learning and its relationship to the PAG mechanisms described here are unclear. Lesions of the lateral habenula do not impair the acquisition of fear conditioning [109]. Moreover, the lateral habenula does not project to the PAG, instead projecting extensively to the ventral mesencephalon [110].

Box 3. The midbrain periaqueductal gray

The PAG is organised into four columns located dorsomedial (dm), dorsolateral (dl), lateral (l), and ventrolateral (vl) to the cerebral aqueduct [14], bordered ventrally by the dorsal raphe (DR) (**Figure 1**). These columns play distinct roles in behaviour and sensory processing and have distinct afferent and efferent connections with other brain regions [60 – 62]. The dorsal columns (dmPAG and dlPAG) control active behavioural coping responses (e.g., escape), whereas ventral columns (lPAG and vlPAG) control passive behavioural coping responses to stressors and threats (e.g., immobility, freezing) [60 – 62]. The PAG has long been implicated in fear and anxiety. Stimulation of the PAG in rats and cats elicits defensive behavioural responses [61] and focal electrical stimulation of PAG in humans generates intense feelings of anxiety [95]. Human neuroimaging studies report increased BOLD signals in the PAG during fear expression, which is maximal at high levels of threat imminence [96, 97]. vlPAG, in particular, receives direct projections from the CeA, notably the medial central nucleus (CeAm), and is critical for expression of conditioned fear responses, including freezing, vocalisation, and conditioned analgesia. Thus, one account of PAG function during Pavlovian fear conditioning emphasises its role in fear response expression, with its columnar organisation subserving defensive response switching or response selection as a function of the imminence of danger [98].

Box 4. Outstanding Questions

- What pathways carry US teaching signal information to the LA? Understanding the neural circuits for predictive fear learning requires an understanding of the pathways obligatory for conveying US related information, including US teaching signals, to the amygdala. Because the PAG does not project directly to LA, these pathways are likely to be indirect, but are currently poorly understood.
- What is the relationship between the role of the PAG in fear expression and fear teaching signals? Learning theories have long divorced the factors determining associative learning from those determining performance of the conditioned response. The PAG is an important part of the neural circuitry controlling the expression of conditioned fear responses and also the circuitry providing prediction error modulated US-teaching signals during Pavlovian fear

conditioning. There is some evidence that these functions are pharmacologically dissociable [28, 29] and there is electrophysiological evidence for distinct populations of CeA-driven PAG neurons [68]. However, more work is needed to understand this relationship.

- Are aversive positive and negative prediction errors coded separately in the fear circuit? A positive prediction error signal for aversive events has been isolated in the PAG and LA but a negative prediction error signal for fear is yet to be unequivocally demonstrated. Understanding the neuronal origin of aversive negative prediction error, whether this signal instructs plasticity in the same or different circuits to aversive positive prediction error, and the role of PAG in these processes, remains key to understanding the fear prediction circuit. It is also critical for understanding fear extinction because aversive negative prediction error is the trigger for extinction learning.
- **What is the role of dopamine in fear learning?** The firing of DA neurons during appetitive learning tasks conforms to the assumptions of associative learning theory. Some DA neurons respond to aversive USs and CSs predictive of such USs [97, 98]. There is also evidence for DA involvement in determining variations in US effectiveness during fear conditioning [104]. However, the function of DA in predictive fear learning, and its relationship to the mechanisms described here, remains enigmatic.
- How are CS associability signals encoded during fear learning? CSs vary in their reliability as signals or predictors of danger. These variations promote changes in CS processing that determine the allocation of attention to the CS and hence its associability with the US. Such changes fall outside the scope of traditional US focussed learning rules such as the Temporal-Difference and Rescorla-Wagner learning rules. A second class of learning rules was developed to accommodate them [15, 109]. The neural pathways for variations in CS processing have been identified in Pavlovian appetitive conditioning [30, 31, 110] but have received little empirical attention in Pavlovian fear conditioning. To date the most compelling evidence is that variations in the effectiveness of the CS as a signal for danger are determined by the nucleus accumbens [100], but how these relate to amygdala mechanisms for fear learning remains unknown.

Acknowledgments

This work was supported, in part, by Australian Research Council grants (DP0343808; DP0877430), National Health and Medical Research Council grants (1003058), and an Australian Research Council QEII Fellowship (DP0877430) to GPM, as well as the National Institutes of Health (R01 MH073700-01) to HTB.

References

1. Davis M. The role of the amygdala in fear and anxiety. *Ann. Rev. Neurosci.* 1992; 15:353–75. [PubMed: 1575447]

2. LeDoux JE. Emotion circuits in the brain. *Ann. Rev. Neurosci.* 2000; 23:155–84. [PubMed: 10845062]
3. Maren S, Quirk GJ. Neuronal signalling of fear memory. *Nat. Rev. Neurosci.* 2004; 5:844–852. [PubMed: 15496862]
4. Sigurdsson T, et al. Long-term potentiation in the amygdala: a cellular mechanism of fear learning and memory. *Neuropharmacology.* 2007; 52:215–227. [PubMed: 16919687]
5. Medina JF, et al. Parallels between cerebellum- and amygdala-dependent conditioning. *Nat. Rev. Neurosci.* 2002; 3:122–131. [PubMed: 11836520]
6. Fanselow MS, Poulos AM. The neuroscience of mammalian associative learning. *Annu Rev Psychol.* 2005; 56:207–234. [PubMed: 15709934]
7. Phelps EA, LeDoux JE. Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron.* 2005; 48:175–187. [PubMed: 16242399]
8. Davis M, et al. Effects of D-cycloserine on extinction: translation from preclinical to clinical work. *Biol. Psychiatry.* 2006; 60:369–375. [PubMed: 16919524]
9. Rosenkranz JA, Grace AA. Dopamine-mediated modulation of odour-evoked amygdala potentials during pavlovian conditioning. *Nature.* 2002; 417:282–287. [PubMed: 12015602]
10. Johansen JP, et al. Optical activation of lateral amygdala pyramidal cells instructs associative fear learning. *Proc. Natl. Acad. Sci. U S A.* 2010; 107:12692–12697. [PubMed: 20615999]
11. Rescorla, RA.; Wagner, AR. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement.. In: Black, AH.; Prokasy, WF., editors. *Classical Conditioning II: Current research and theory.* Appleton Century Crofts; 1972. p. 64-99.
12. Sutton RG, Barto AG. Toward a modern theory of adaptive networks: Expectation and prediction. *Psychol. Rev.* 1981; 88:135–170. [PubMed: 7291377]
13. Kamin, LJ. 'Attention-like' processes in Classical conditioning.. In: Jones, MR., editor. *Miami symposium on the prediction of behavior: aversive stimulation.* University of Miami Press; 1968. p. 9-33.
14. Bradfield L, McNally GP. Unblocking in Pavlovian fear conditioning. *J. Exp. Psychol. Anim. Behav. Process.* 2008; 34:256–265. [PubMed: 18426308]
15. Pearce JM, Hall G. A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psych. Rev.* 1980; 87:532–552.
16. Kim JJ, et al. Inhibitory cerebello-olivary projections and blocking effect in classical conditioning. *Science.* 1998; 279:570–573. [PubMed: 9438852]
17. Gutfreund Y, et al. Gated visual input to the central auditory system. *Science.* 2002; 297:1556–1559. [PubMed: 12202831]
18. Waelti P, et al. Dopamine responses comply with basic assumptions of formal learning theory. *Nature.* 2001; 412:43–48. [PubMed: 11452299]
19. Schultz W, Dickinson A. Neuronal coding of prediction errors. *Ann. Rev. Neurosci.* 2000; 23:473–500. [PubMed: 10845072]
20. Belova MA, et al. Expectation modulates neural responses to pleasant and aversive stimuli in primate amygdala. *Neuron.* 2007; 55:970–984. [PubMed: 17880899]
21. Johansen JP, et al. Neural substrates for expectation-modulated fear learning in the amygdala and periaqueductal gray. *Nat. Neurosci.* 2010; 13:979–986. [PubMed: 20601946]
22. Di Scala G, et al. Evidence of Pavlovian conditioned fear following electrical stimulation of the periaqueductal grey in the rat. *Physiol Behav.* 1987; 40:55–63. [PubMed: 3615655]
23. Fanselow MS, Bolles RC. Naloxone and shock-elicited freezing in the rat. *J. Comp. Physiol. Psychol.* 1979; 93:736–744. [PubMed: 479405]
24. Young SL, Fanselow MS. Associative regulation of Pavlovian fear conditioning: unconditional stimulus intensity, incentive shifts, and latent inhibition. *J. Exp. Psychol. Anim. Behav. Process.* 2002; 18:400–413. [PubMed: 1402696]
25. McNally GP, et al. Blocking, unblocking, and overexpectation: a role for opioid receptors in the regulation of Pavlovian association formation. *Behav. Neurosci.* 2004; 118:111–120. [PubMed: 14979787]

26. Zelikowsky M, Fanselow MS. Opioid regulation of Pavlovian overshadowing in fear conditioning. *Behav. Neurosci.* 2010; 124:510–519. [PubMed: 20695650]
27. Eippert F, et al. Blockade of endogenous opioid neurotransmission enhances acquisition of conditioned fear in humans. *J. Neurosci.* 2008; 28:5465–5472. [PubMed: 18495880]
28. McNally GP, Cole S. Opioid receptors in the midbrain periaqueductal gray regulate prediction errors during pavlovian fear conditioning. *Behav. Neurosci.* 2006; 120:313–23. [PubMed: 16719696]
29. Cole S, McNally GP. Opioid receptors mediate direct predictive fear learning: evidence from one-trial blocking. *Learn. Mem.* 2007; 14:229–235. [PubMed: 17404385]
30. Calu DJ, et al. Neural correlates of variations in event processing during learning in central nucleus of amygdala. *Neuron.* 2010; 68:991–1001. [PubMed: 21145010]
31. Roesch MR, et al. Neural correlates of variations in event processing during learning in basolateral amygdala. *J. Neurosci.* 2010; 30:2464–2471. [PubMed: 20164330]
32. Cole S, McNally GP. Temporal-difference prediction errors and Pavlovian fear conditioning: role of NMDA and opioid receptors. *Behav. Neurosci.* 2007; 121:1043–1052. [PubMed: 17907835]
33. Pearce, JM.; Mackintosh, NJ. Two theories of attention: a review and possible integration.. In: Mitchell, CJ.; LePelley, ME., editors. *Attention and Associative Learning*. Oxford: 2010. p. 11-40.
34. LePelley ME. The role of associative history in models of associative learning: A selective review and hybrid model. *Q. J. Exp. Psychol.* 2004; 57B:193–243.
35. Quirk GJ, Mueller D. Neural mechanisms of extinction learning and retrieval. *Neuropsychopharmacology.* 2008; 33:56–72. [PubMed: 17882236]
36. Ehrlich I, et al. Amygdala inhibitory circuits and the control of fear memory. *Neuron.* 2009; 62:757–671. [PubMed: 19555645]
37. Pape HC, Pare D. Plastic synaptic networks of the amygdala for the acquisition, expression, and extinction of conditioned fear. *Physiol. Rev.* 2010; 90:419–463. [PubMed: 20393190]
38. McNally GP, Westbrook RF. Opioid receptors regulate the extinction of Pavlovian fear conditioning. *Behav. Neurosci.* 2003; 117:1292–1301. [PubMed: 14674848]
39. Kim JH, Richardson R. The effect of the mu-opioid receptor antagonist naloxone on extinction of conditioned fear in the developing rat. *Learn. Mem.* 2009; 16:161–166. [PubMed: 19223599]
40. McNally GP, et al. Opioid receptors in the midbrain periaqueductal gray regulate extinction of pavlovian fear conditioning. *J. Neurosci.* 2004; 24:6912–6919. [PubMed: 15295026]
41. McNally GP, et al. The midbrain periaqueductal gray and fear extinction: opioid receptor subtype and roles of cyclic AMP, protein kinase A, and mitogen activated protein kinase. *Behav. Neurosci.* 2005; 119:1023–1033. [PubMed: 16187830]
42. Parsons RG, et al. Regulation of extinction-related plasticity by opioid receptors in the ventrolateral periaqueductal gray matter. *Front. Behav. Neurosci.* 2010; 4:ii, 44.
43. McNally GP. Facilitation of fear extinction by midbrain periaqueductal gray infusions of RB101(S), an inhibitor of enkephalin-degrading enzymes. *Behav. Neurosci.* 2005; 119:1672–1677. [PubMed: 16420170]
44. Hugues S, et al. Prefrontal infusion of PD098059 immediately after fear extinction training blocks extinction-associated prefrontal synaptic plasticity and decreases prefrontal ERK2 phosphorylation. *Synapse.* 2006; 60:280–287. [PubMed: 16786530]
45. Herry C, et al. Extinction of auditory fear conditioning requires MAPK/ERK activation in the basolateral amygdala. *Eur. J. Neurosci.* 2006; 24:261–269. [PubMed: 16882022]
46. Davis M. Role of NMDA receptors and MAP kinase in the amygdala in extinction of fear: clinical implications for exposure therapy. *Eur. J. Neurosci.* 2002; 16:395–398. [PubMed: 12193180]
47. Kozak AT, et al. Naltrexone renders one-session exposure therapy less effective: a controlled pilot study. *J. Anx. Dis.* 2007; 21:142–152.
48. Merluzzi TV, et al. Opioid antagonist impedes exposure. *J. Consult. Clin. Psych.* 1991; 59:425–430.
49. Egan KJ. Endogenous opiate system and systematic desensitization. *J. Consult. Clin. Psych.* 1988; 56:287–291.

50. Bryant RA, et al. A study of the protective function of acute morphine administration on subsequent posttraumatic stress disorder. *Biol Psychiat.* 2009; 65:438–440. [PubMed: 19058787]
51. Bolles RC, Fanselow MS. A perceptual defensive recuperative model of fear and pain. *Behav. Brain Sci.* 1980; 3:291–323.
52. Hawkins RD, Kandel ER. Is there a cell biological alphabet for simple forms of learning? *Psychol. Rev.* 1984; 91:375–391. [PubMed: 6089240]
53. McLaren, IPL. The computational unit as an assembly of neurones: an implementation of an error correcting learning algorithm.. In: Durbin, R., et al., editors. *The Computing Neuron*. Addison-Wesley; 1989. p. 160-78.
54. Fanselow MS. Pavlovian conditioning, negative feedback, and blocking: mechanisms that regulate association formation. *Neuron.* 1998; 20:625–627. [PubMed: 9581755]
55. McNally GP. The roles of endogenous opioids in fear learning. *Int. J. Comp. Psych.* 2009; 22:153–169.
56. Tracey, DJ. Ascending and descending pathways in the spinal cord.. In: Paxinos, G., editor. *The Rat Nervous System*. 3rd edn. Academic; 2004. p. 149-164.
57. Ciochi S, et al. Encoding of conditioned fear in central amygdala inhibitory circuits. *Nature.* 2010; 468:277–282. [PubMed: 21068837]
58. Haubensak W, et al. Genetic dissection of an amygdala microcircuit that gates conditioned fear. *Nature.* 2010; 468:270–276. [PubMed: 21068836]
59. Duvarci S, et al. Central Amygdala Activity during Fear Conditioning. *J. Neurosci.* 2011; 31:289–294. [PubMed: 21209214]
60. Keay, KA.; Bandler, R. The periaqueductal gray.. In: Paxinos, G., editor. *The Rat Nervous System*. 3rd edn. Academic; 2004. p. 243-257.
61. Carrive P. The periaqueductal gray and defensive behavior: functional representation and neuronal organization. *Behav. Brain Res.* 1993; 58:27–47. [PubMed: 8136048]
62. Bandler R, et al. Integration of somatic and autonomic reactions within the midbrain periaqueductal grey: viscerotopic, somatotopic and functional organization. *Prog. Brain Res.* 1991; 87:269–305. [PubMed: 1678189]
63. Gonçalves L, et al. Prefrontal afferents to the dorsal raphe nucleus in the rat. *Brain Res. Bull.* 2009; 78:240–247. [PubMed: 19103268]
64. Corcoran KA, Quirk GJ. Activity in prelimbic cortex is necessary for the expression of learned, but not innate, fears. *J Neurosci.* 2007; 27:840–844. [PubMed: 17251424]
65. Keay KA, Bandler R. Parallel circuits mediating distinct emotional coping reactions to different types of stress. *Neurosci. Biobehav. Rev.* 2001; 25:669–678. [PubMed: 11801292]
66. Amat J, et al. Medial prefrontal cortex determines how stressor controllability affects behavior and dorsal raphe nucleus. *Nat. Neurosci.* 2005; 8:365–371. [PubMed: 15696163]
67. Amat J, et al. Previous experience with behavioral control over stress blocks the behavioral and dorsal raphe nucleus activating effects of later uncontrollable stress: role of the ventral medial prefrontal cortex. *J. Neurosci.* 2006; 26:13264–12272. [PubMed: 17182776]
68. da Costa Gomez TM, Behbehani MM. An electrophysiological characterization of the projection from the central nucleus of the amygdala to the periaqueductal gray of the rat: the role of opioid receptors. *Brain Res.* 1995; 689:21–31. [PubMed: 8528703]
69. Tershner SA, Helmstetter FJ. Antinociception produced by mu opioid receptor activation in the amygdala is partly dependent on activation of mu opioid and neurotensin receptors in the ventral periaqueductal gray. *Brain Res.* 2000; 865:17–26. [PubMed: 10814729]
70. Helmstetter FJ, Tershner SA. Lesions of the periaqueductal gray and rostral ventromedial medulla disrupt antinociceptive but not cardiovascular aversive conditional responses. *J. Neurosci.* 1994; 14:7099–7108. [PubMed: 7965101]
71. Heinricher MM. Evidence for two classes of nociceptive modulating neurons in the periaqueductal gray. *J Neurosci.* 1987; 7:271–278. et a. [PubMed: 3806198]
72. Krout KE, Loewy AD. Periaqueductal gray matter projections to midline and intralaminar thalamic nuclei of the rat. *J. Comp. Neurol.* 2000; 424:111–141. [PubMed: 10888743]

73. Berendse HW, Groenewegen HJ. Restricted cortical termination fields of the midline and intralaminar thalamic nuclei in the rat. *Neuroscience*. 1991; 42:73–102. [PubMed: 1713657]
74. Groenewegen HJ, Berendse HW. The specificity of the ‘nonspecific’ midline and intralaminar thalamic nuclei. *Trends in Neurosci*. 1994; 17:52–57.
75. Su HS, Bentivoglio M. Thalamic midline cell populations projecting to the nucleus accumbens, amygdala, and hippocampus in the rat. *J. Comp. Neurol*. 1990; 297:582–593. [PubMed: 1696591]
76. Jasmin L, et al. Rostral agranular insular cortex and pain areas of the central nervous system: a tract-tracing study in the rat. *J. Comp. Neurol*. 2004; 468:425–440. [PubMed: 14681935]
77. Price DD. Psychological and neural mechanisms of the affective dimension of pain. *Science*. 2000; 288:1769–1772. [PubMed: 10846154]
78. Jensen J, et al. Separate brain regions code for salience vs. valence during reward prediction in humans. *Hum. Brain Mapp*. 2007; 28:294–302. [PubMed: 16779798]
79. Seymour B, et al. Temporal difference models describe higher-order learning in humans. *Nature*. 2004; 429:664–667. [PubMed: 15190354]
80. Seymour B, et al. Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. *Nat. Neurosci*. 2005; 8:1234–1240. [PubMed: 16116445]
81. Spoormaker VI, et al. The neural correlates of negative prediction error signaling in human fear conditioning. *Neuroimage*. 2011; 54:2250–2256. [PubMed: 20869454]
82. Dunsmoor JE, et al. Neural correlates of unconditioned response diminution during Pavlovian conditioning. *Neuroimage*. 2008; 40:811–817. [PubMed: 18203622]
83. Furlong TM, et al. The role of prefrontal cortex in predictive fear learning. *Behav. Neurosci*. 2010; 124:574–586. [PubMed: 20939658]
84. Shi C, Davis M. Pain pathways involved in fear conditioning measured with fear-potentiated startle: lesion studies. *J. Neurosci*. 1999; 19:420–430. [PubMed: 9870970]
85. Lanuza E, et al. Unconditioned stimulus pathways to the amygdala: effects of posterior thalamic and cortical lesions on fear conditioning. *Neuroscience*. 2004; 125:305–315. [PubMed: 15062974]
86. Brunzell DH, Kim JJ. Fear conditioning to tone but not context is attenuated by lesions of insular cortex and posterior extension of intralaminar complex. *Behav. Neurosci*. 2001; 115:365–375. [PubMed: 11345961]
87. Morgan MA, LeDoux JE. Differential contribution of dorsal and ventral medial prefrontal cortex to the acquisition and extinction of conditioned fear in rats. *Behav. Neurosci*. 1995; 109:681–688. [PubMed: 7576212]
88. Lacroix L, et al. Differential role of the medial and lateral prefrontal cortices in fear and anxiety. *Behav. Neurosci*. 2000; 114:1119–1130. [PubMed: 11142644]
89. Johansen JP, Fields HL. Glutamatergic activation of anterior cingulate cortex produces an aversive teaching signal. *Nat. Neurosci*. 2004; 7:398–403. [PubMed: 15004562]
90. Tang J, et al. Pavlovian fear memory induced by activation in the anterior cingulate cortex. *Mol. Pain*. 2005; 1:6. [PubMed: 15813993]
91. Bissière S, et al. The rostral anterior cingulate cortex modulates the efficiency of amygdala-dependent fear learning. *Biol. Psychiatry*. 2008; 63:821–831. [PubMed: 18155183]
92. Gao YJ, et al. Contributions of the anterior cingulate cortex and amygdala to pain- and fear-conditioned place avoidance in rats. *Pain*. 2004; 110:343–353. [PubMed: 15275785]
93. Zhang F, et al. Circuit-breakers: optical technologies for probing neural signals and systems. *Nat. Rev. Neurosci*. 2007; 8:577–581. [PubMed: 17643087]
94. Desai M, et al. Mapping brain networks in awake mice using combined optical control and fMRI. *J. Neurophysiol*. 2010 doi: 10.1152/jn.00828.2010.
95. Richardson DE, Akil, H. H. Pain reduction by electrical brain stimulation in man. 1: acute administration in periaqueductal and periventricular sites. *J. Neurosurg*. 1977; 47:178–183. [PubMed: 327030]
96. Mobbs D, et al. From threat to fear: the neural organization of defensive fear systems in humans. *J. Neurosci*. 2009; 29:12236–12243. [PubMed: 19793982]
97. Mobbs D, et al. When fear is near: threat imminence elicits prefrontal-periaqueductal gray shifts in humans. *Science*. 2007; 317:1079–1083. [PubMed: 17717184]

98. Fanselow, MS. The midbrain periaqueductal gray as a coordinator of action in response to fear and anxiety.. In: Depaulis, A.; Bandler, R., editors. *The Midbrain Periaqueductal Gray Matter: Functional, Anatomical and Neurochemical Organization*. Plenum; 1991. p. 151-173.
99. Matsumoto M, Hikosaka O. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*. 2009; 459:837–841. [PubMed: 19448610]
100. Brischoux F, et al. Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proc Natl Acad Sci U S A*. 2009; 106:4894–4899. [PubMed: 19261850]
101. Iordanova MD, et al. Dopamine activity in the nucleus accumbens modulates blocking in fear conditioning. *Eur. J. Neurosci*. 2006; 24:3265–3270. [PubMed: 17156387]
102. Iordanova MD, et al. Opioid receptors in the nucleus accumbens regulate attentional learning in the blocking paradigm. *J. Neurosci*. 2006; 26:4036–4045. [PubMed: 16611820]
103. Guarraci FA, et al. Amygdaloid D1 dopamine receptor involvement in Pavlovian fear conditioning. *Brain Res*. 1999; 827:28–40. [PubMed: 10320690]
104. Greba Q, Kokkinidis L. Peripheral and intraamygdalar administration of the dopamine D1 receptor antagonist SCH 23390 blocks fear-potentiated startle but not shock reactivity or the shock sensitization of acoustic startle. *Behav. Neurosci*. 2000; 114:262–272. [PubMed: 10832788]
105. Nader K, LeDoux JE. Inhibition of the mesoamygdala dopaminergic pathway impairs the retrieval of conditioned fear associations. *Behav. Neurosci*. 1999; 113:891–901. [PubMed: 10571473]
106. Iordanova MD. Dopamine transmission in the amygdala modulates surprise in an aversive blocking paradigm. *Behav. Neurosci*. 2010; 124:780–788. [PubMed: 21133534]
107. Matsumoto MM, Hikosaka O. Representation of negative motivational value in the primate lateral habenula. *Nat. Neurosci*. 2009; 12:77–84. [PubMed: 19043410]
108. Konorski, J. *Integrative activity of the brain*. University of Chicago Press; 1967.
109. Heldt SA, Ressler KJ. Lesions of the habenula produce stress- and dopamine-dependent alterations in prepulse inhibition and locomotion. *Brain Res*. 2006; 1073-1074:229–239. [PubMed: 16442084]
110. Kim U. Topographic commissural and descending projections of the habenula in the rat. *J. Comp. Neurol*. 2009; 513:173–187. [PubMed: 19123238]
111. Mackintosh NJ. A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychol. Rev*. 1975; 82:276–298.
112. Holland, PC.; Maddux, J-M. Brain systems for attention in associative learning.. In: Mitchell, CJ.; LePelley, ME., editors. *Attention and Associative Learning*. Oxford: 2010. p. 305-350.

Figure 1.

Fear learning and US evoked responding in rat LA and PAG neurons is attenuated when the shock US is expected and depends on MOR in the vIPAG. (a) Rats were trained in a two-stage fear conditioning paradigm. In the first stage (‘Stage I’), animals were trained with CS(A) US pairings (A+) over three days and the percentage of freezing behavior during the 30-second CS presentations was recorded. During ‘Stage II’, animals received either vehicle (Sal) or CTAP (MOR antagonist) administered into the vIPAG prior to combined pairings of either CS(AB)-US [in which the US was already predicted by CS(A)] or CS(CD) -US (in which the US was not predicted). During the ‘Test’ stage, behavioral freezing responses to 30 s presentations of CS(B) (blue) and CS(D) (black) were assessed drug free and without the shock US. Blocking of fear learning (i.e. reduced fear learning) to CSB was observed in animals which had previously received intra-vIPAG saline, as exemplified by a lower freezing during the 30 s presentations of CS(B) compared with CS(D) (‘Saline’). However, the blocking effect was abolished in animals which had previously received intra-vIPAG CTAP (‘CTAP’). (b) & (c) Rats were trained in a Pavlovian fear conditioning involving electrophysiological recordings of LA and PAG neurons during CS and US presentations. The US-evoked neural response was significantly inhibited in both the LA and PAG when it

is predicted by a well-trained CS. Population averaged (Y-axis) peri-event time histograms showing inhibition of US-evoked responding in (b) LA and (c) PAG neurons when the US is predicted (Pred, blue) by a previously trained CS compared with when it is presented unpredictably (Unpred, black). Time during the US presentation (individual 2 ms shock pulses over 2 seconds) is shown on the X-axes with individual shock pulses indicated by red hash marks. Note that statistical analyses compared averaged firing rates during the US period in the Pred and Unpred conditions. Panel A is reproduced, with permission, from [28]. Panel B is reproduced, with permission, from [21].

Figure 2.

Working model of proposed neural circuitry that is involved in teaching signal processing during fear conditioning. For simplicity, ascending projections are shown in the left hemisphere only and descending projections in the right hemisphere only. The putative US pathway conveying information about the actual shock US (λ) is shown in black. The putative CS pathway conveying information about the expected outcome ($-\Sigma V$) is shown in red. The putative prediction error modulated teaching signal ($\lambda - \Sigma V$) is shown in green. Shock USs (λ) activate spinal and trigeminal dorsal horn neurons which project to the PAG and from there through midline and intralaminar thalamus to the dmPFC and LA to produce depolarization of LA pyramidal cells. During fear conditioning, weaker auditory thalamic and cortical CS afferent inputs to LA pyramidal neurons (broken black lines) are strengthened ($-\Sigma V$) when they are co-active with US-evoked depolarization of the same cells. Following conditioning, CS inputs to LA activate projection neurons to the CeA, which send inhibitory projections to the vIPAG ($-\Sigma V$) that: 1) produce freezing (possibly by relieving

PAG output neurons from tonic inhibition); and, 2) inhibit shock-US responsive neurons in the PAG to attenuate US processing. Disinhibited output from the PAG may also be relayed to the rostroventromedial medulla to inhibit shock US processing at the level of the dorsal horn. Thus, the ascending projection from the dorsal horn to the PAG may also be a prediction error modulated teaching signal. Pathways for CS-mediated inhibition would serve to inhibit US-evoked depolarization of LA neurons when the US is predicted (such as during blocking) thereby limiting associative plasticity of CS inputs onto the same cells.

Figure 3.

Expectation modulates thalamic and cortical responses to aversive USs. (a) Human participants received CS presentations followed by a loud noise unconditioned stimulus (UCS) on 100% ('CS100') or 50% ('CS50') of trials. The unexpected UCS elicited a greater BOLD response in the anterior cingulate and dorsolateral PFC as compared to the expected UCS, as revealed by fMRI recordings of subjects during the fear conditioning experiment (AUC: area under curve for BOLD signal in anterior cingulate on CS50 and CS100 trials [left bottom]). The magnitude of the US-elicited fMRI signal (AUC) was negatively correlated with the self-report of UCS expectancy in both dorsolateral PFC and anterior cingulate (anterior cingulate shown in right bottom). (b) Time course data for fMRI signal in thalamus. fMRI signal was similar during the 10 s CS period (CS onset and duration indicated by grey background) for both CS50 (solid line) and CS100 (dashed line) but responses to the loud noise UCS (UCS onset and duration indicated by black bar) were greater on the CS50 (i.e. UCS unexpected) than CS100 (i.e. UCS expected) trials ((c & d) Rats received auditory CS and shock US presentations. These were either expected ('Expected') due to prior CS – US pairings or unexpected due to the absence of such prior pairings ('Unexpected'). Unexpected CS – US presentations produced significantly greater neuronal activation (as measured by c-Fos expression) in midline thalamic neurons

identified as projecting to the dmPFC, compared to expected presentations or CS alone presentations ('Control'). Black dots in (c) indicate c-Fos positive nuclei; brown dots indicate neurons retrogradely labeled from the dmPFC with the neural tracer cholera toxin B (CTb); blue dots indicate dual labeled CTb/c-Fos positive neurons. Panels A and B are reproduced, with permission, from [82]. Panels C and D are modified with permission from [83].

Box 1 Figure I.

Schematic diagram depicting the main areas of the amygdala and its involvement in fear learning and memory.

Box 3 Figure I.

Schematic diagram depicting the main anatomical subregions of the PAG.