View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by UCL Discovery

The translational neural circuitry of anxiety

Oliver Robinson*[1], Alexandra Pike[1], Brian Cornwell[2], Christian Grillon[3]

[1]Neuroscience and Mental Health group, Institute of Cognitive Neuroscience & Research Department of Clinical, Educational and Health Psychology, University College London, London, UK

[2]Centre for Mental Health, Swinburne University of Technology, Melbourne, Australia

[3]National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland, USA

*Corresponding author

Wordcount: 5378 words

References: 60 in main text, 97 supplemental

Abstract

Anxiety is an adaptive response that promotes harm avoidance; but at the same time excessive anxiety constitutes the most common psychiatric complaint. Moreover, current treatments for anxiety – both psychological and pharmacological - hover at around 50% recovery rates. Improving treatment outcomes is nevertheless difficult, in part because contemporary interventions were developed without an understanding of the underlying neurobiological mechanisms that they modulate. Recent advances in experimental models of anxiety in humans, such as threat of unpredictable shock, have, however, enabled us to start translating the wealth of mechanistic animal work on defensive behavior into humans. In this paper we discuss the distinction between fear and anxiety, before reviewing translational research on the neural circuitry of anxiety in animal models and how it relates to human neuroimaging studies across both healthy and clinical populations. We highlight the roles of subcortical regions (and their subunits) such as the bed nucleus of the stria terminalis, the amgydala, and the hippocampus, as well as their connectivity to cortical regions such as dorsal medial and lateral prefrontal/cingulate cortex and insula in maintaining anxiety responding. We discuss how this circuitry

might be modulated by current treatments before finally highlighting areas for future research that might ultimately improve treatment outcomes for this common and debilitating transdiagnostic symptom.

1Introduction

Fear and anxiety are adaptive, defensive reactions to threat across species. However, excessive fear or anxiety can interfere with quality of life. Indeed, anxiety disorders are the most prevalent psychiatric disorders and excessive anxiety is implicated in most psychiatric disorders, as well as a number of other medical and neurological conditions. Anxiety is thus accompanied by a high financial cost[1]. Compounding this, response rates to first-line pharmacological and psychological treatments are less than 50%[2]: most patients fail to respond to the first treatment that targets their anxiety. Other review papers focus on the neural circuitry of fear[3–5]; this review focuses on our understanding of the underlying neurobiology of anxiety, as a construct not only more closely related to anxiety disorders such as generalised anxiety disorder (GAD), but also highly relevant transdiagnostically to other psychiatric and neurological disorders. In this review, which is intended as a broad narrative introduction for those new to the field, we argue that understanding this neurobiology, as well as the features that differentiate *adaptive* from *pathological* anxiety, is key to identifying pathological mechanisms and treatment targets.

Fear and anxiety

Whilst fear and anxiety share many subjective and physiological symptoms, they can be differentiated based on behavioral profiles determined by certainty[6,7], which can be further subdivided into the contingency, temporal precision, and spatial precision of the threat:

- In fear, the danger is imminent, unambiguous and mobilizes the organism to take immediate action. Fear is above all a rapid behavioral response that leads to active avoidance (e.g., fight-or-flight), or other automatic responses such as freezing in prey animals or piloerection (goosebumps). Pathological fear is seen in specific phobias, which are characterized by a marked fear of specific objects.
- In anxiety, the threat is more diffuse and uncertain. Anxiety is a lasting state of apprehension of potential future threats accompanied by negative affect, autonomic symptoms, worry, increased vigilance, and passive avoidance[8]. Excessive anxiety symptoms can be found in GAD and panic disorder (though panic attacks themselves may be better characterized by models of acute fear, panic disorder, including apprehension of subsequent attacks, is considered to be better modelled by anxiety)[8].

There is a body of research indicating that fear and anxiety are tractable and are associated with distinct pathologies[6,7,9]. Early research into fear and anxiety found that there was a double dissociation between neural structures relating to threats that were phasic (fear) and sustained (anxiety)[6], which led to a theoretical model in which anxiety and fear are putatively separate processes. Although the true neurobiological picture is likely to be more nuanced than this[10,11], this distinction is also reflected in the RDoC matrix, which distinguishes between acute threat (fear) and potential threat (anxiety)[12,13] as well as in the DSM-5 where specific phobias are defined as "cued by the presence or anticipation of *specific* objects or situations", while generalized anxiety disorder, by contrast, is defined as "Excessive anxiety and worry (apprehensive expectation) … about *a number of* events or activities"[14].

The widest array of research to date involves Pavlovian cued *fear conditioning* in rodents[15]. Unfortunately, while fear conditioning is a useful model of fear, it is insensitive to drugs that are anxiolytic in humans[16–19] and is thus a poor model of anxiety disorders such as GAD[20,21]. Comparatively little is known about anxiety, especially its many human-specific cognitive-affective features. Indeed, many animal models of anxiety, such as the elevated-plus maze, have few analogs in humans[22] (although see Biederman et al. (2017)[23] and Bach et al. (2014)[24]) , and of course the impact of psychological therapies cannot be studied in animals. However, innovative approaches to study anxiety experimentally in humans have recently been developed. This paper reviews this emerging literature and suggests a model of its neural underpinnings.

The phenomenology of pathological anxiety

Following Freud who distinguished chronic anxiety from anxiety (panic) attack, clinicians have long recognized that anxiety is not a unitary phenomenon[8]. This non-unitary view of anxiety is reflected in the DSM-5, which identifies several anxiety disorders characterized by shared "features of excessive fear and anxiety", including panic disorder (PD), generalized anxiety disorder (GAD), social anxiety disorder (SAD), and simple phobia (SP). Other disorders within the DSM also have anxiety as a core symptom, such as obsessive-compulsive disorders and addiction disorders. Additionally, a number of neurological disorders feature elevated anxiety, including variants of dementia such as frontotemporal dementia, vascular dementia, and Alzheimer's disease[25], along with Parkinson's disease[26] and traumatic brain injury[27,28].

However, despite symptom heterogeneity, we do not at present have clear objective markers that can differentiate between disorders which feature anxiety, and there is, moreover, strong symptom overlap. All share common enduring behavioral, cognitive, and physiological characteristics, potentially arising from impairments in transdiagnostic features, as seen in the RDoC[12] domain of negative valence systems[13] which highlights exaggerated or problematic responses to 'potential threat (anxiety)'. In this paper, based on the assumption that similar phenomenological presentations of 'anxiety' reflect true underlying neurobiological similarities, we discuss the shared neural circuitry which may underlie sustained anxiety symptoms.

Translational neuroscience of anxiety

Early work in fear conditioning in animal models highlighted a key role of two amygdala nuclei, the basolateral amygdala (BLA) and the central nucleus of the amygdala (CeA) in anxiety. The BLA integrates sensory information from the environment and, via its projections, excites the CeA. The amygdala subsequently triggers defensive responses, via efferent projections to regions such as the stria terminalis, the hippocampus, the ventral striatum, the orbitofrontal cortex, the periaqueductal grey, and the hypothalamus[29]. While the amygdala is important for fear conditioning, its direct role in maintaining sustained anxiety symptoms has been more difficult to establish. Lesions of the amygdala do not reduce defensive responses in models of anxiety such as the elevated plus-maze[30] and the anxiolytic benzodiazepine does not act via the amygdala[31].

This contrasts with another structure tightly coupled with the CeA, the bed nucleus of the stria terminalis (BNST), which does appear to be preferentially involved in maintaining sustained anxiety[6,7,32]. The BNST is a part of the "extended amygdala"[33], which is well-situated to regulate defensive responses such as anxiety via its GABA (gamma-aminobutyric acid)-ergic projections to various limbic, hindbrain and cortical structures[6,7,32] .

Consideration of the conditions that determine the involvement of the BNST in defensive responses emphasizes the role of temporal unpredictability of the threats[34] and the sustained duration of the response[7,35]. Early evidence of a differentiation between fear and anxiety in the BNST came from studies using the *startle reflex*. 'Fear-potentiated startle' refers to the increased startle reflex amplitude in the presence of a short-duration threat, while 'anxiety-potentiated startle' to the increased startle amplitude during a long-duration unpredictable threat. A series of studies by Davis and collaborators

established a double dissociation between the CeA and BNST; lesions of the CeA abolish fear-, but not anxiety-potentiated startle, while lesions of the BNST suppress anxiety-, but not fear-potentiated startle[6,7]. According to Davis' group, anxiety is thus maintained by activation of corticotrophin receptors in the BNST[7].

More recent studies have, however, provided evidence that there is more nuance to the role of the BNST in anxiety. This is perhaps unsurprising given that the BNST is small but heterogeneous, with up to eighteen functionally distinct sub-regions[36]. Some BNST lesions can *up-regulate* anxiety[37], while optogenetic stimulation of discrete BNST sub-regions can *down-regulate* anxiety[38,39]. Different efferent connections of the BNST also control different features of anxiety[38,39], suggesting that distinct BNST sub-regions dynamically control different aspects of defensive behavior[38–44] (see Table 1). Of note, the story is almost certainly the same with regards to the role of the amygdala in fear, which is gradually being broken down into component micro-circuits[45,46].

Table 1: Studies informing our emerging understanding of the functions of the subregions of the BNST. Note that the majority of work has been done in rodents, due to the lack of spatial resolution of MRI prior to the advent of 7T MRI scanning.

Region of BNST	Function	Species
Right anterior	Activated during threat>safe contrast in healthy control participants[40]	Human
Anteroventral	Paraventricular nucleus excitation, promotes corticosterone secretion[41]	Rat
Posterior	Inhibits stress-induced paraventricular nucleus excitation, adrenocorticotropic hormone release, and plasma corticosterone responses[41]	Rat
Anterolateral (encompassing oval nucleus)	Attenuates stress-induced reductions in weight gain at end of 7 day stress period: mediating consequences of repeated stress?[42]	Rat
Anterodorsal BNST	Social behaviour: namely defensive and reproductive behaviour[43]; promotes anxiolytic behavioour e.g. dereased risk-avoidance, decreased respiratory rate, positive conditioning valence[38]	Rat

Oval nucleus of the BNST	Anxiogenic (perhaps by inhibiting anterodorsal BNST)[38]	Rat
Ventral BNST	Innervates VTA, stimulating glutamatergic projections results in aversive/anxiogenic phenotypes, stimulating GABAergic perojections results in rewarding/anxiolytic phenotypes by inhibition of VTA GABAergic neurons[39]; noradrenergic neurons may control freezing behaviour[44]	Rat

Moving into the cortex, the medial prefrontal cortex (mPFC) is implicated in attention and affective information processing[47] and hence also involved in responses to threat. The rodent mPFC is comprised of several sub-regions including the infralimbic (IL) and prelimbic cortex (PL), which are densely interconnected with the amygdala[48] and play opposing roles in defensive behaviors[49], with the IL inhibiting the amygdala (and defensive behaviors) and the PL exciting the amygdala. Empirical evidence of a role for the PL in sustained anxiety has been established using a wide range of paradigms including the elevated plus maze and the open field test[50], context conditioning[51], and fear conditioning with long-duration conditioned stimuli (CS)[52]. Electrophysiological recordings in rodents show that PL neurons maintain persistent firing that correlate with freezing throughout the duration of sustained threat[51,52] and that PL firing persists during trace fear conditioning[53], a paradigm during which the US is delivered after an empty interval following the end of the CS. These results suggest PL involvement in both defensive behaviors and the neural representation of threat (i.e., the unconditioned stimulus or US).

The hippocampus also conveys contextual information about environmental threat[54] to the PL[55]. Specifically, brain structures communicate via synchronized activity both in local networks[56] and over longer distances[57]. This can be observed in theta oscillations, which have been linked to anxiety in rodents and humans[58] and display synchronized activity between the ventral hippocampus and PL in anxiogenic contexts[54]. Additionally, research using operant conflict tasks has suggested that the hippocampus has a key role in decision-making in situations where there is a conflict between approaching rewards and avoiding punishments, a scenario which induces anxiety across-species[59].

Interoceptive information from visceral changes may also be conveyed, via the anterior insula, to the PL[60]. Specifically, robust visceral changes (heart rate, gastrointestinal, blood pressure) caused by

anxiety may generate a feedback loop between the PL and anterior insula that contributes to the maintenance of anxiety. Threat representation in the PL could then influence anxiety responses via direct[s1] or indirect input to the BNST through the basolateral amygdala[60].

Finally, PFC regions are also involved in anxiety via their role in working memory (WM). In rodents, WM deficits are associated with increased anxiety[s2]. Primates have larger prefrontal cortices with additional medial and lateral dissociations. In monkeys, as in humans, WM relies in part on the dorsolateral PFC (dlPFC)[s3], a structure that is also implicated in anxiety and stress. Maternal separation stress in young monkeys activates the right dlPFc, but deactivates the left dlPFC[s4]. In young monkeys, as in children, heightened reaction to novelty and potential threat characterizes an anxious temperament disposition, which is a well-established risk for the development of anxiety[s5]. Critically, such anxious disposition in monkeys is associated with dlPFC malfunction[s6].

A possible neural model of anxiety that emerges from these studies in animals is that environmental signals from the ventral hippocampus and interoceptive signals from the anterior insula help maintain threat representation in the PL, which is then used to guide defensive behaviors via the BNST and the amygdala, and may be under the control of the dlPFC (see Figure 1).

[Insert Figure 1 here]

The neuroscience of human anxiety and its pathology

In recent years innovative translational experimental models that attempt to mimic and quantify sustained anxiety in humans have emerged. These include darkness[s7], low concentration (7.5%) of CO2 inhalation[s8], and vigilant threat monitoring[s9] (see Table 2). However, one widely used anxiety model in humans relies on long-duration threat of aversive stimuli such as shock (i.e., threat of shock or context conditioning[s10]), where subjects are informed that shocks may be delivered unpredictably.

Other paradigms investigating fear (e.g. 'cued fear conditioning'[s11]) may seem at first sight to be investigating similar constructs. However, these paradigms are generally much more precise as to the temporal (and spatial) precision of the shock association (notably, further research is needed to ascertain the precise temporal/spatial conditions under which anxiety-related - rather than fear-related - circuitry is activated).

Table 2: Paradigms used for studying anxiety (as contrasted here for utility with fear). These paradigms often include aversive stimuli such as shocks, but may also be used with other aversive stimuli such as aversive noises/pictures or blasts of cold air. The commonality among all these paradigms is that they all evoke a sustained anxiety state.

Paradigm	Description
Threat-of-shock	Threat-of-shock paradigms instruct participants that they may be in one of two conditions: 'safe from shock' (during which no shocks are received) or 'at risk of shock' (during which shocks may be received). Shocks may occur during the 'at risk' blocks but with a low probability, and generally are scheduled independently of task performance. Other tasks may be performed with blocks of these contexts, to investigate the effects of this sustained threat on other cognitive processes. This is often measured using the eyeblink reflex, in a procedure known as 'startle'. Greater anxiety is thought to correspond to the augmentation of the eyeblink reflex.
Low dose (7.5%) CO_2	Inhaling air with > 7.5% CO_2 concentration over a period of around 20 minutes increases both subjective experience of and physiological symptoms associated with anxiety
Darkness	As humans are a diurnal species, this task elicits anxiety and potentiates anxious responses (e.g. startle) by contrasting lighted conditions with complete darkness
Vigilant threat monitoring	Participants observe a line fluctuating on the screen which they are told reflects their own physiological levels of 'anxiety' (though in fact this line is generated by experimenters), and when this line passes a threshold they will receive a shock. They are instructed to remain calm and avoid accumulating shocks.
Context conditioning	During normal conditioning procedures, conditioning occurs either to the stimulus (cued conditioning) and/or to the context (context conditioning). The context here refers to features of the environment that are present during the conditioning procedure. Context conditioning is separate to the typical 'cued conditioning' used in fear learning paradigms due to the decreased temporal precision of the association between the context and the unconditioned stimulus.
Long-duration CS	This paradigm is similar to typical fear conditioning, in that a predictive stimulus (CS) is paired with an unconditioned aversive stimulus (US), except that the duration of presentation of the CS is longer, typically upwards of 30s.
Approach-avoidance conflict	This type of paradigm, a translation of the animal paradigm of operant conflict, involves a choice or conflict between approaching rewards and avoiding punishments and has been argued to be anxiogenic.

One key advantage of the startle probe methodology, discussed in the animal section above, is that it can also be employed in humans. Healthy human participants display heightened startle sensitivity during unpredictable threat relative to baseline, but, importantly, this startle sensitivity is further elevated in those with disorders that feature anxiety[7,s12]. Specifically, multiple studies have now shown exaggerated anxiety-potentiated (but not fear-potentiated) startle during unpredictable threat in panic disorder[s13–s16] and in PTSD [s17–s19]. GAD does not follow quite the same pattern – it is associated

with increased startle overall during a threatening experimental environment. Exaggerated anxiety-potentiated startle may therefore constitute a risk factor or early biomarker for developing disorders including heightened anxiety[s18,s19].

Findings from unpredictable threat of shock studies in humans have implicated many of the same behavioral and neural responses that underpin (adaptive) anxiety in animals. In turn, these responses have also been implicated in pathological anxiety in humans. For instance, early functional magnetic resonance imaging (fMRI) studies demonstrated amygdala involvement in processing both predictable[s20] and unpredictable threat[9]. Activity in the amygdala was also shown to be elevated in those with social phobia relative to controls[s21,s22] and it was argued that this elevated response was critical for the development of the excessive anxiety.

Subsequent work has, however, tempered these conclusions by demonstrating that the amygdala is sensitive to appetitive as well as aversive stimuli[s23]. Moreover, consistent with animal data, many studies fail to report selective amygdala activation during sustained anxiety symptoms[s24,s25]. It has therefore been argued that perhaps a key role of the amygdala is in goal-directed cognitive processing and/or behavior towards relevant/salient information[s26]. Given the clear asymmetries between appetitive and aversive stimuli value (i.e. the consequences are generally worse if one misses a threat than a reward), it may be that prior work associating the amygdala with anxiety reflects a more fundamental role of the amygdala that happens to be correlated with symptoms of anxiety (for example elevated harm avoidance in anxiety promoting the relevance/salience of threats) rather than any *selective* role in anxiety.

Animal models also highlight the role of the prefrontal cortex in anxiety. In humans, elevated fMRI activity in dorsal regions of the PFC (dorsomedial PFC; dmPFC and dorsal anterior cingulate; dACC) has been associated with both unpredictable threat processing and pathological anxiety[s27]. It has been argued, moreover, that it is specifically the rostral part of the dmPFC that drives conscious threat appraisal and worry[s28]. In healthy humans, dorsal ACC/dmPFC activation has been associated with a reduced ability to extinguish fear responding[s29], and patients with PTSD show increased activation in the dACC[s30]. As with the amygdala, again, it should be noted that the dACC/dmPFC is responsible for a wide range of functions including goal directed behaviour, vestibular function, social responding and interoception[s31]. Critically, it is hyperactive in most psychiatric disorders[s32] and, as with the amygdala, this may be due to a more generic role in anticipating emotional stimuli[s33] and hence a key role in

appraising and expressing behavioural responses to the level of environmental threat[s34]. In other words, hyperactivity in psychiatric disorders probably reflects this region's more fundamental role - such directing cognitive processing and/or behaviour towards relevant/salient information[s32] – that is important for harm avoidance.

Perhaps unsurprisingly given the links between both amygdala and dACC/dmPFC activity and anxiety-relevant symptoms, *connectivity* between these regions has also been implicated in the pathophysiology of anxiety. Specifically, functional imaging studies show that connectivity between the amygdala and the mPFC increases when healthy individuals are exposed to threat of unpredictable shock, and that the strength of this connectivity is greater in individuals with higher dispositional anxiety[s35–s37]. This extends to the pathological state, such that in those with social anxiety or GAD, this same circuitry is elevated without overt anxiety induction[s38]. As such, the same circuitry that drives heightened attention towards threats under adaptive anxiety is also critical in pathological anxiety. To this end, it has been argued that this circuitry represents a human (functional) homologue of the rodent PL-amygdala/BNST circuit highlighted above[s30,s34–s37]. If so, the key role that such circuitry plays in behavioral response to salient aversive information could underlie the harm avoidant negative bias towards threats in pathological anxiety.

As also highlighted in the animal work above, the hippocampus is involved in anxiety, due, perhaps, to its key role in contextual learning/memory[s39] and prospection[s40] or, alternatively, a role in avoidance[24,s41,s42]. Consistent with functional differentiation along the longitudinal axis of the rodent hippocampus[s43], theta activity (2-8 Hz) from anterior hippocampus (or ventral in rodents) correlates with anxiety level in humans[58]; and at the same time theta from posterior hippocampus correlates with spatial memory performance in a simulated human version of the rodent Morris water maze task[s44]. Notably, this theta-based coupling between hippocampus and mPFC scales-up with increased threat probability[s45] and patients with PTSD exhibit aberrant activity and connectivity involving the hippocampus in the resting state[s46]. This theta coupling evidence comes from magnetoencephalography (MEG): the evidence is more mixed in fMRI, with both hypo- and hyper-activation of the hippocampus reported in patients[s47]. Finally, threat of shock also modulates memory performance; improving contextual learning under threatening conditions[s48,s49]; which may be one mechanism by which the hippocampus maintains traumatic memories in disorders which feature elevated anxiety such as PTSD[s50]. Avoidance, on the other hand, can be studied using approach-avoidance conflict tasks (a variant on operant conflict tasks outlined above). Such tasks have been translated from non-human primate

work[s51,s52] into humans[24,s41,s42,s53]. This paradigm incorporates a range of tasks that set up a conflict between the inherent bias to avoid (learned or prepotent) negative outcomes and an approach response. This is thought to either be anxiogenic in itself[24,s41,s42] , or, at the very least, to elicit avoidance responses[s54] which are a core feature of anxiety disorders. Critically, such tasks implicate the hippocampus [24,s41,s42] and this conflict has been shown to be exacerbated, leading to increased avoidance, in humans during induced anxiety[s55], and in pathological anxiety[s56].

The insula also plays a complex role in anxiety in humans. Together with the ACC and mPFC, the insula is a part of a network that detects, interprets, and reacts to internal bodily signals[s57]. Interoceptive signals are thought to be integrated in the insula following a *posterior-mid-anterior* pattern, with processing in the anterior insula producing conscious awareness of the information[s58].The anterior insula has been suggested to make a key contribution to the anticipation and emotional experience of aversive stimuli, and, via the ACC, to the allocation of attention and initiation of appropriate action[s57]. Experimental psychopathology studies show that the anterior insula is broadly involved in anticipation of aversive events, but more specifically during anticipation of unpredictable compared to predictable threat[s59] and sustained versus transient anticipation[9]. Thus, the anterior insula is one of the structures that reliably maintains sustained activation during experimentally-induced anxiety[9]. It is also hyperactive in pathological anxiety, including panic disorder[s60] and GAD[s61], during sustained threat. In this instance, pathophysiology is not associated with chronic activation of the anterior insula but rather with heightened response during anxiety induction, possibly reflecting a feeling of lack of control[s62] as well as autonomic and emotional distress during threat[s23].

While animal studies have implicated the BNST in anxiety for some time now, it is only recently that progress in the spatial resolution of MRI has led to comprehensive exploration of this small structure in humans[s63,s64]. Older studies found activity in regions overlapping the BNST during induced anxiety - i.e. unpredictable threats[9,s9,s24] - as distinct from amygdala activity during predictable threats (i.e. fear)[9,s65]. In addition, BNST activation during unpredictable shock correlates positively with the magnitude of autonomic arousal[s62]. Consistent with findings from startle studies, heightened sensitivity to unpredictable threat in PTSD and panic disorders is associated with elevated sustained BNST activation, making this structure a promising biomarker of disease and treatment targeting[s60,s66]. These studies also identified several structures that are coactivated with the BNST, including the dmPFC, ventrolateral; PFC, dorsolateral PFC, and anterior insula, attesting to the complexity of a putative

‘anxiety network’. These results suggest an important role for the BNST in mediating the hyperarousal and hypervigilance symptoms of pathological anxiety.

Additionally, high resolution imaging using resting-state fMRI has revealed connectivity of the BNST in humans to many of the other regions discussed above[s64,s67]. Critically, anxiety induced by threat of shock reveals reduced BNST connectivity with the ventromedial prefrontal cortex (vmPFC) and ACC[s63]. These are early findings, and research exploring clinical anxiety with high resolution imaging is still in its infancy, but given the highlighted role of the BNST in the animal literature, clinical research targeting the BNST, its microcircuits, and its functional connectivity offers great promise in the development of novel therapeutic interventions to treat pathological anxiety[32].

Finally, as in non-human primates, the dlPFC is involved in anxiety regulation and dispositional anxiety in humans[s68], potentially because of a role in emotion regulation and attention control. In healthy subjects, the dlPFC is activated during anxiety induction procedures and the strength of this activation is negatively correlated with anxiety, indicating that poor dlPFC activation is associated with less anxiety down-regulation (although this is speculative)[s69]. Moreover, the dlPFC plays a role in explicit emotion regulation (for a review across all types or regulation strategies see[s70]). When subjects perform a cognitive task that engages the dlPFC (e.g. a working memory task – a ‘distraction’ form of emotional regulation), dlPFC activation concomitantly reduces anxiety induced by unpredictable shock, via top-down control exerted on the dACC and ventrolateral PFC[s71]. The dlPFC is therefore a key structure for healthy functioning. When the dlPFC is activated to keep task goals in mind, dlPFC engagement also suppresses emotional interference and alleviates anxiety. It is now well-established that the dlPFC is hypo-activated in those with psychiatric disorders featuring anxiety during cognitive tasks, emotion regulation studies, and during anxiety induction procedures[s32,s72].

To summarize, research shows overlapping neural circuitry in response to anxiety between humans and animals. Specifically, connectivity between the hippocampus, BNST amygdala and medial prefrontal/cingulate cortex may contribute to a putative ‘anxiety network’ that may also be (down-) regulated by dorsolateral PFC. Notably, this network of regions bears some similarity to the ‘salience network’, which incorporates regions such as the dorsal anterior cingulate and insular regions[s73]. Nevertheless, many of the regions (especially the PFC) are not clear translational homologues across species and, perhaps more importantly, human work highlights that many of these patterns of activity are not unique to anxiety and may instead reflect more fundamental cognitive processes, such as

salience processing (in accordance with the key identity of many of those nodes within the salience network); which happen to constitute a key facet of anxiety.

The role of treatment on anxiety circuitry

Characterizing the behavioral effects of current treatments for pathological anxiety as well as their underlying neural mechanisms is crucial for the development of new treatments and the improvement of existing ones.

Cornwell et al. (2017) recently reported MEG evidence that hypervigilant responding under threat of unpredictable shocks[s74] is reduced by the benzodiazepine alprazolam[s75] and that this is driven by increased feedback signaling from ventrolateral prefrontal to sensory cortices. Although not directly implicating a role for any of the structures reviewed above, these data point to possible alternative avenues to examine the efficacy of novel anxiolytic treatments.

When given chronically to healthy individuals, the selective serotonin reuptake inhibitor (SSRI) citalopram selectively reduces anxiety-potentiated startle (but not fear-potentiated startle) to unpredictable threat[s76]. Interestingly, acute citalopram *increases* anxiety-potentiated startle[s77], an effect consistent with the clinical observations of transiently increased anxiety during initial SSRIs treatment[s78]. Both the anxiolytic effects of chronic SSRI administration and the anxiogenic effects of acute SSRI administration have been replicated in rodents[s79] and the current view is that these effects involve the BNST[7] and result from interactions between serotonin and corticotropin-releasing factors[7,s79].

Acute tryptophan depletion (ATD) is a dietary manipulation that can be used to temporarily reduce levels of serotonin. ATD is associated with *increased* engagement of the dmPFC-amygala circuit that has been shown to be hyper-engaged by both threat of shock[s35] and pathological anxiety[s38] in humans. SSRIs may therefore work, at least in part, by elevating synaptic serotonin availability and hence *reducing* engagement of this dorsal prefrontal- amygdala circuit[s80]. Indeed, it has been argued that the magnitude of SSRI response is predicted by *greater* pretreatment reactivity to threats in pregenual ACC and lesser reactivity in the amygdala[s81]. In anxious patients, SSRIs have also been shown to attenuate connectivity between the BNST and limbic and paralimbic structures[s82]. One hypothesis, therefore, is that SSRIs work by attenuating putative 'anxiety network' hyperactivity[s83], but this remains to be tested.

Another broad class of treatment for anxiety disorders is psychological interventions. The goal of the most common psychological intervention, cognitive behavioral therapy (CBT), is to attenuate negative mood states through cognitive reappraisal and emotion regulation strategies[s84]. Medial prefrontal and amygdala activity have been argued to predict treatment response to CBT in anxiety[s85], perhaps through modulation of circuitry between the medial prefrontal cortex and amygdala. Indeed a recent meta-analysis suggested that the most robust predictors in response to therapy were significant post-therapy decreases in anterior cingulate/paracingulate gyrus, inferior frontal gyrus and insula activity[s86].

One problem with cross-sectional observational studies, however, is that it is unclear if neural circuitry change is specifically due to the intervention or rather reflects a generic change in symptoms (i.e. it reflects symptom change, not mechanistic change). One way to study this is to explicitly modulate basic processes targeted by CBT in the absence of symptom changes. To this end, in healthy individuals, simple attentional instruction can alter the engagement of threat of shock induced dorsomedial PFC-amygdala circuitry[s37]. Specifically, asking individuals to re-appraise emotional stimuli as neutral *dampens down* activity in the circuitry thought to drive heightened response to threats in anxiety. Ongoing work is exploring whether this is a mechanism by which CBT for anxiety works.

Exposure is another important psychological intervention, with similar efficacy to CBT[s87]. Exposure is primarily targeted at fear responding (e.g., phobias), but can also be used to reduce anxiety. A more detailed review on findings from extinction paradigms – extinction is thought to be the mechanism by which exposure therapy works - in both humans and animals can be found in Milad and Quirk (2012)[s88], but studies have implicated some of the same circuitry – activity was reduced in the amygdala, insula and anterior cingulate, but increased in the dlPFC[s89] following extinction.

In sum, preliminary work suggests that current treatments for anxiety may be effective through modulation of the translational circuitry outlined above. This work is in its infancy, but ultimately, a mechanistic understanding of treatment response may eventually enable improvement of existing treatments, better targeting of existing treatments to patients who will respond, and provide targets for the design of novel treatments: hence resulting in increased recovery rates for anxiety disorders.

Future perspectives

1. One key problem with the current categorical disease classifications for anxiety disorders is that of heterogeneity within diagnostic categories and overlapping symptoms across disorders. Specifically, anxiety seems to be an obvious candidate for an RDoC-based approach, with the RDoC matrix highlighting several constructs (acute threat, potential threat (anxiety), and sustained threat) that map onto the distinction made in this paper between anxiety and fear. This distinction may map on to different types of vulnerability: perhaps one associated with predictable threat (i.e., fear) and the other with unpredictable threat (i.e., anxiety). Of relevance, this distinction appears in line with data of factor analytic studies showing a distinction for 'fear disorder' and 'anxious-misery' in the anxiety disorders[s90]. Large-scale studies will be necessary to obtain sufficient power to determine the extent to which anxiety circuitry is compromised (in similar, and in different ways) across psychiatric and neurological disorders, which could improve classification, treatment planning, and pave the way for precision medicine approaches. Anxious patients with heightened sensitivity to unpredictable threat (e.g., panic disorder[s13]) for example may particularly benefit from treatment that down-regulate the anxiety circuit (e.g., SSRIs[s76]). This approach can be applied more broadly: those with elevated sensitivity to unpredictable threat, whatever their psychiatric/neurological diagnosis, may benefit from the same treatments.
2. The BNST has long been overlooked. It is a small but functionally complex structure comprised of at least 18 sub-regions[36] involved, amongst other functions, in opposing anxiolytic and anxiogenic circuits. Basic research in rodents on intrinsic circuits of the BNST and how these circuits are impacted by stress hormones and neurotransmitters will be crucial to increase our knowledge of normal and pathological anxiety and to develop treatment[s91]. For example, the actions of neuropeptide corticotropin-releasing factor (CRF) on the BNST in relation to sustained anxiety provided a strong rationale for the therapeutic development of CRF1 antagonists to treat mood and anxiety disorders. However, these have failed in human models[s92] and in clinical trials[s93]. More studies on the local and more distal connectivity of the BNST will be necessary to understand why CRF antagonists are anxiolytic in animal models but not in humans. On this note, a question remains about how best to translate our understanding of such subunits into humans. Improved, higher resolution, fMRI techniques will undoubtedly help, but fMRI signal is

still several steps away from the underlying neural activity. Work with specific PET ligands and even direct electrical recording from those with clinical implants may be needed to ultimately corroborate translational patterns.

3. Relatedly, direct translation across humans and animals is difficult. Animal models are essential to advance mechanistic understanding of behaviors, but their limitations must be acknowledged. Pathological anxiety is, above all, a disorder of feelings supported by conscious and unconscious experiences, and animal models rely on overt behaviour and physiological measures, rather than cognitions or subjective experiences. Additionally, many processes, e.g. the psychological symptoms targeted by CBT, cannot be directly studied in animal models, and many regions – especially cortical areas – are likely not direct functional homologues across animal models (if they exist at all in animals with smaller cortices). This is of particular concern because many animal models of treatment have failed to translate to humans. To overcome this, future work may seek to use models of anxiety in healthy humans alongside the exact same models in animals as putative anxiolytic drug screens to ensure successful translation. Furthermore, cross-fertilization may be important where roadblocks in treatment development occur: theories and biological insights from one type of research may produce advances in the other.
4. Beyond pharmacology, neurostimulation research is starting to bear exciting results. Deep brain stimulation of the BNST reduces anxiety in a rodent model and in humans[s94]. Repetitive transcranial magnetic stimulation (rTMS) is a less invasive neurostimulation technique. As low-frequency oscillations within cortical networks decrease cortical excitability[s92], targeting the dmPFC or dLPFC-amygdala coupling with low (1Hz) (rTMS) could down-regulate activity in this circuit and ultimately reduce negative affect. rTMS of the dlPFC is approved by the FDA as a second-line treatment for depression, and research is currently undergoing to examine its effectiveness in anxiety[s95].
5. Future research should also focus on the emergence of pathological anxiety, ideally taking a developmental perspective. There are a number of difficulties inherent in developmental work in anxiety, with the most important being the ethical issues with exposing children and adolescents to aversive events such as shock[s96]. However, researchers have recently successfully collected data on fear-potentiated startle in adolescents using alternative aversive stimuli such

as cold air blasts or aversive screams[s96,s97,]. These are promising steps towards creating paradigms that can elucidate the developmental basis anxiety.

Concluding remarks

Our understanding of anxiety circuitry has grown considerably thanks to experimental psychopathology models exploring the impact of unpredictable threats. Consistent with animal models, fear and anxiety in humans may be mediated by dissociable, although partly overlapping, neural mechanisms. Structures implicated in anxiety but not fear include the BNST, hippocampus, dmPFC, insula, and dlPFC. Fear- and anxiety-related defensive responses are mediated by the CeA and BNST, respectively.

From a clinical and treatment perspective, there is overlap in the circuitry implicated in disorders featuring anxiety and in the mechanism of action of reference anxiolytics. Pathophysiological mechanisms are found in the same neural structures that respond adaptively to anxiety induction procedures in healthy humans. Neural dysfunction can take two forms, chronic activation (inappropriate activation in the absence of anxiety induction challenge, e.g., heightened amygdala-dmPFC connectivity) or exaggerated activation in response to an unpredictable threat (i.e., insula).

Despite this progress, much remains to be learned. With advances in technologies in basic (optogenetic, molecular biology, transgenic and knockout mice) and clinical research (high spatial resolution of fMRI, better statistical tools), the focus should be on improving knowledge of local microcircuits and neural oscillations among distant regions that supports behavior. Clinically, research on fear and anxiety circuitry might be used to create an evidence-based nosology[s18]. Ultimately, improved, personalized, and new treatment strategies will be difficult to develop without a better understanding of the underlying mechanisms of anxiety; the work reviewed here constitutes a step forward, but a precise mechanistic understanding is still far off.

Acknowledgements and Disclosures:

Contributorship: O.J.R, A.C.P, B.C and C.G are all responsible for drafting and editing this manuscript.

Funding: This work was supported by the personal fellowships MR/K024280/1 and MR/R020817/1 from the Medical Research Council to OJR and the Intramural Research Program of the National Institute of

Mental Health, project number ZIAMH002798 (clinical protocol 02-M-0321 (NCT00047853), 01-M-0185 (NCT00026559)) to CG. OJR has completed consultancy work for IESO digital health and Brainbow and is running an Investigator Initiated Trial with Lundbeck. He holds an MRC Industrial Collaboration Award with Cambridge Cognition.

Competing interests: The authors report no competing interests.

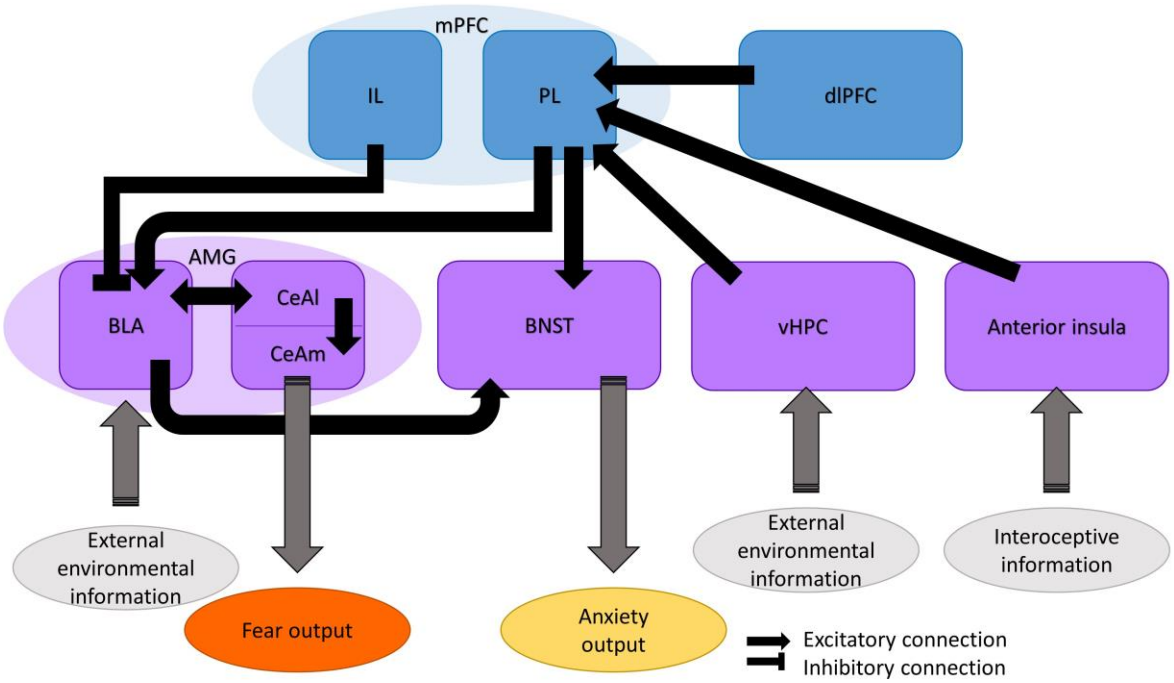

Figure 1 caption

Figure 1: A simplified diagram of fear/anxiety circuitry, derived from animal research. IL = infralimbic cortex; PL = paralimbic cortex, vHPC = ventral hippocampus, BLA = basolateral amygdala, CeA = central nucleus of the amygdala, AMG = amygdala. A more detailed diagram of part of this circuitry can be found in Calhoon & Tye, 2015. Note that the dlPFC is included on the basis of research in non-human primates: rodents may not have a region homologous with the dlPFC.

References

1. Greenberg, P. E. *et al.* The economic burden of depression in the United States: how did it change between 1990 and 2000? *The Journal of clinical psychiatry* **64**, 1465–1475 (2003).
2. Baldwin, D. S., Huusom, A. K. T. & Mæhlum, E. Escitalopram and paroxetine in the treatment of generalised anxiety disorder: randomised, placebo-controlled, double-blind study. *The British Journal of Psychiatry* **189**, 264–272 (2006).
3. Kim, J. J. & Jung, M. W. Neural circuits and mechanisms involved in Pavlovian fear conditioning: A critical review. *Neuroscience & Biobehavioral Reviews* **30**, 188–202 (2006).
4. Tovote, P., Fadok, J. P. & Lüthi, A. Neuronal circuits for fear and anxiety. *Nature Reviews Neuroscience* **16**, 317–331 (2015).
5. Fullana, M. A. *et al.* Neural signatures of human fear conditioning: an updated and extended meta-analysis of fMRI studies. *Molecular Psychiatry* **21**, 500–508 (2016).
6. Davis, M. Are different parts of the extended amygdala involved in fear versus anxiety? *Biological Psychiatry* **44**, 1239–1247 (1998).
7. Davis, M., Walker, D. L., Miles, L. & Grillon, C. Phasic vs sustained fear in rats and humans: role of the extended amygdala in fear vs. anxiety. *Neuropsychopharmacol* **35**, 105–135 (2010).
8. Barlow, D. H. Unraveling the mysteries of anxiety and its disorders from the perspective of emotion theory. *American Psychologist* **55**, 1247–1263 (2000).
9. Alvarez, R. P., Chen, G., Bodurka, J., Kaplan, R. & Grillon, C. Phasic and sustained fear in humans elicits distinct patterns of brain activity. *Neuroimage* **55**, 389–400 (2011).
10. Lázaro-Muñoz, G., LeDoux, J. E. & Cain, C. K. Sidman Instrumental Avoidance Initially Depends on Lateral and Basal Amygdala and Is Constrained by Central Amygdala-Mediated Pavlovian Processes. *Biological Psychiatry* **67**, 1120–1127 (2010).
11. Evans, D. A. *et al.* A synaptic threshold mechanism for computing escape decisions. *Nature* **558**, 590–594 (2018).
12. Insel, T. *et al.* Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. (2010).
13. *NIMH Research Domain Criteria (RDoC) project Negative Valence Systems: Workshop Proceedings*. (2011).
14. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (DSM-5)*. (American Psychiatric Publishing, 2013).
15. Phelps, E. A. & Ledoux. J.E. Contributions of the amygdala to emotion processing: From animal models to human behavior. *Neuron* **48**, 175–187 (2005).
16. Grillon, C. *et al.* The benzodiazepine alprazolam dissociates contextual fear from cued fear in humans as assessed by fear-potentiated startle. *Biol Psychiatry* **60**, 760–766 (2006).
17. Baas, J. *et al.* Benzodiazepines have no effect on fear-potentiated startle in humans. *Psychopharmacology* **161**, 233–247 (2002).

18. Moberg, C. A. & Curtin, J. J. Alcohol selectively reduces anxiety but not fear: Startle response during unpredictable versus predictable threat. *Journal of Abnormal Psychology* **118**, 335–347 (2009).

19. Graham, S. J. *et al.* Effects of lorazepam on fear-potentiated startle responses in man. *Journal of Psychopharmacology* **19**, 249–258 (2005).

20. Blanchard, R. J., Yudko, E. B., Rodgers, R. J. & Blanchard, D. C. Defense system psychopharmacology: an ethological approach to the pharmacology of fear and anxiety. *Behavioural and Brain Research* **58**, 155_165 (1993).

21. Maier, S. F. & Watkins, L. R. Stressor controllability, anxiety, and serotonin. *Cognitive therapy and research* **22**, 595–613 (1998).

22. Walz, N., Mühlberger, A. & Pauli, P. A human open field test reveals thigmotaxis related to agoraphobic fear. *Biol Psychiatry* (2016).

23. Biedermann, S. V. *et al.* An elevated plus-maze in mixed reality for studying human anxiety-related behavior. *BMC Biology* **15**, (2017).

24. Bach, D. R. *et al.* Human Hippocampus Arbitrates Approach-Avoidance Conflict. *Current Biology* **24**, 541–547 (2014).

25. Porter, V. R. *et al.* Frequency and Characteristics of Anxiety Among Patients with Alzheimer's Disease and Related Dementias. *The Journal of Neuropsychiatry and Clinical Neurosciences* **15**, 180–186 (2003).

26. Broen, M. P. G., Narayen, N. E., Kuijf, M. L., Dissanayaka, N. N. W. & Leentjens, A. F. G. Prevalence of anxiety in Parkinson's disease: A systematic review and meta-analysis: Prevalence of Anxiety in PD. *Movement Disorders* **31**, 1125–1133 (2016).

27. Scholten, A. C. *et al.* Prevalence of and Risk Factors for Anxiety and Depressive Disorders after Traumatic Brain Injury: A Systematic Review. *Journal of Neurotrauma* **33**, 1969–1994 (2016).

28. Hart, T. *et al.* Prevalence, Risk Factors, and Correlates of Anxiety at 1 Year After Moderate to Severe Traumatic Brain Injury. *Archives of Physical Medicine and Rehabilitation* **97**, 701–707 (2016).

29. LeDoux, J. E. Emotion circuits in the brain. *Ann Rev Neurosc* **23**, 155–184 (2000).

30. Treit, D., Pesold, C. & Rotzinger, S. Dissociating the anti-fear effects of septal and amygdaloid lesions using two pharmacologically validated models of rat anxiety. *Behavioral Neuroscience* **107**, 770–785 (1993).

31. Blanchard, D. C., Griebel, G. & Blanchard, R. J. Mouse defensive behaviors: pharmacological and behavioral assays for anxiety and panic. *Neurosci Biobehav Rev* **25**, 205–18 (2001).

32. Lebow, M. A. & Chen, A. Overshadowed by the amygdala: the bed nucleus of the stria terminalis emerges as key to psychiatric disorders. *Mol Psychiatry* (2016). doi:10.1038/mp.2016.1

33. Heimer, L., Harlan, R. E., Alheid, G. F., Garcia, M. M. & de Olmos, J. J. Substantia innominata: a notion which impedes clinical-anatomical correlations in neuropsychiatric disorders. *Neuroscience* **76**, 957–1006 (1997).

34. Goode, T. D. & Maren, S. Role of the bed nucleus of the stria terminalis in aversive learning and memory. *Learn Mem* **24**, 480–491 (2017).

35. Hammack, S., Todd, T., Kocho Schellenberg, M. & Bouton, M. Role of the bed nucleus of the stria terminalis in the acquisition of contextual fear at long or short context-shock intervals. *Behavioral Neuroscience* **129**, 673–678 (2015).

36. Bota, M., Sporns, O. & Swanson, L. W. Neuroinformatics analysis of molecular expression patterns and neuron populations in gray matter regions: The rat BST as a rich exemplar. *Brain Research* **1450**, 174–193 (2012).

37. Duvarci, S., Bauer, E. & Pare, D. The bed nucleus of the stria terminalis mediates inter-individual variations in anxiety and fear. *The Journal of Neuroscience* **29**, 10357–61 (2009).

38. Kim, S.-Y. *et al.* Diverging neural pathways assemble a behavioural state from separable features in anxiety. *Nature* **496**, 219–223 (2013).

39. Jennings, J. H. *et al.* Distinct extended amygdala circuits for divergent motivational states. *Nature* **496**, 224–228 (2013).

40. Avery, S. N., Clauss, J. A. & Blackford, J. U. The Human BNST: Functional Role in Anxiety and Addiction. *Neuropsychopharmacology* **41**, 126–141 (2016).

41. Choi, D. C. *et al.* Bed Nucleus of the Stria Terminalis Subregions Differentially Regulate Hypothalamic-Pituitary-Adrenal Axis Activity: Implications for the Integration of Limbic Inputs. *Journal of Neuroscience* **27**, 2025–2034 (2007).

42. Roman, C. W. *et al.* Excitotoxic lesions of the bed nucleus of the stria terminalis (BNST) attenuate the effects of repeated stress on weight gain: Evidence for the recruitment of BNST activity by repeated, but not acute, stress. *Behavioural Brain Research* **227**, 300–304 (2012).

43. Dong, H.-W. & Swanson, L. W. Projections from bed nuclei of the stria terminalis, posterior division: Implications for cerebral hemisphere regulation of defensive and reproductive behaviors. *The Journal of Comparative Neurology* **471**, 396–433 (2004).

44. Fendt, M. Noradrenaline Transmission within the Ventral Bed Nucleus of the Stria Terminalis Is Critical for Fear Behavior Induced by Trimethylthiazoline, a Component of Fox Odor. *Journal of Neuroscience* **25**, 5998–6004 (2005).

45. Ciocchi, S. *et al.* Encoding of conditioned fear in central amygdala inhibitory circuits. *Nature* **468**, 277–282 (2010).

46. Haubensak, W. *et al.* Genetic dissection of an amygdala microcircuit that gates conditioned fear. *Nature* **468**, 270–276 (2010).

47. Whalen, P. J. *et al.* The emotional counting Stroop paradigm: a functional magnetic resonance imaging probe of the anterior cingulate affective division. *Biol Psychiatry* **44**, 1219–28 (1998).

48. McDonald, A. J. Cortical pathways to the mammalian amygdala. *Prog Neurobiol* **55**, 257–332 (1998).

49. Sotres-Bayon, F. & Quirk, G. J. Prefrontal control of fear: more than just extinction. *Current Opinion in Neurobiology* **20**, 231–235 (2009).

50. Bi, L.-L. *et al.* Enhanced excitability in the infralimbic cortex produces anxiety-like behaviors. *Neuropharmacology* **72**, 148–156 (2013).

51. Baeg, E. H. *et al.* Fast Spiking and Regular Spiking Neural Correlates of Fear Conditioning in the Medial Prefrontal Cortex of the Rat. *Cerebral Cortex* **11**, 441–451 (2001).

52. Burgos-Robles, A., Vidal-Gonzalez, I. & Quirk, G. J. Sustained conditioned responses in prelimbic prefrontal neurons are correlated with fear expression and extinction failure. *The Journal of Neuroscience* **29**, 8474–82 (2009).

53. Gilmartin, M. R. & McEchron, M. D. Single neurons in the medial prefrontal cortex of the rat exhibit tonic and phasic coding during trace fear conditioning. *Behavioral Neuroscience* **119**, 1496–1510 (2005).

54. Adhikari, A., Topiwala, M. A. & Gordon, J. A. Synchronized Activity between the Ventral Hippocampus and the Medial Prefrontal Cortex during Anxiety. *Neuron* **65**, 257–269 (2010).

55. Swanson, L. W. A direct projection from Ammon's horn to prefrontal cortex in the rat. *Brain Res* **217**, 150–154 (1981).

56. Gray, C. M., Konig, P., Engel, A. K. & Singer, W. Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature* **338**, 334–337 (1989).

57. von Stein, A. & Sarnthein, J. Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization. *International Journal of Psychophysiology* **38**, 301–313 (2000).

58. Cornwell, B. R., Arkin, N., Overstreet, C., Carver, F. W. & Grillon, C. Distinct contributions of human hippocampal theta to spatial cognition and anxiety. *Hippocampus* **22**, 1849–1859 (2012).

59. Ito, R. & Lee, A. C. H. The role of the hippocampus in approach-avoidance conflict decision-making: Evidence from rodent and human studies. *Behavioural Brain Research* **313**, 345–357 (2016).

60. Vertes, R. P. Differential projections of the infralimbic and prelimbic cortex in the rat. *Synapse* **51**, 32–58 (2004).