Dynamic competition between large-scale functional networks differentiates fear conditioning and extinction in humans

Lars Marstaller[1,2], Hana Burianová[1,3], and David C. Reutens[1,2]

[1]Centre for Advanced Imaging, University of Queensland, Brisbane, Australia

[2]ARC Science of Learning Research Centre, University of Queensland, Brisbane, Australia

[3]ARC Centre of Excellence in Cognition and its Disorders, Macquarie University, Sydney, Australia

Correspondence should be addressed to:

Lars Marstaller, PhD
Centre for Advance Imaging
University of Queensland
QLD 4072, AUSTRALIA
l.marstaller@uq.edu.au

Pages: 23

Figures: 5

Abstract

The high evolutionary value of learning when to respond to threats or when to inhibit previously learned associations after changing threat contingencies is reflected in dedicated networks in the animal and human brain. Recent evidence further suggests that adaptive learning may be dependent on the dynamic interaction of meta-stable functional brain networks. However, it is still unclear which functional brain networks compete with each other to facilitate associative learning and how changes in threat contingencies affect this competition. The aim of this study was to assess the dynamic competition between large-scale networks related to associative learning in the human brain by combining a repeated differential conditioning and extinction paradigm with independent component analysis of functional magnetic resonance imaging data. The results (i) identify three task-related networks involved in initial and sustained conditioning as well as extinction, and demonstrate that (ii) the two main networks that underlie sustained conditioning and extinction are anti-correlated with each other and (iii) the dynamic competition between these two networks is modulated in response to changes in associative contingencies. These findings provide novel evidence for the view that dynamic competition between large-scale functional networks differentiates fear conditioning from extinction learning in the healthy brain and suggest that dysfunctional network dynamics might contribute to learning-related neuropsychiatric disorders.

Keywords: associative learning; dynamic connectivity; neuroimaging

Introduction

Learning the predictive value of external stimuli is fundamental to successful behavioural adaptation and is governed by specialized networks in the mammalian brain. Current models of cortical dynamics suggest that the human brain exhibits dynamic changes in connectivity, resulting in a meta-stable system in which a number of brain regions temporarily synchronize their activity to form transiently stable functional networks, whereas competing brain regions transiently desynchronize their activity (Shanahan, 2010; Deco & Jirsa, 2012). The view of cortical dynamics suggests that learning-related disorders, which lead to behavioural maladaptation, such as post-traumatic stress disorder or generalized anxiety, may be associated with disturbed dynamic interactions between such large-scale functional networks (Uhlhaas & Singer, 2012; Calhoun et al., 2014; Kringelbach et al., 2015; Li et al., 2014; Panzeri et al., 2015; Zalesky et al., 2014). Therefore, the key to understanding the pathophysiology of learning-related disorders is the characterization of the structure and dynamics of the functional networks that specifically subserve fear and extinction learning. To date, however, surprisingly little is known about these networks and their interactions in the normal adult human brain.

Evidence from animal studies identifies several neural circuits, involving the amygdala, hippocampus, and ventromedial prefrontal cortex, which are engaged during aversive conditioning and extinction (Fanselow, 1994; LeDoux, 2000; Lang, Davis & Öhman, 2001; Sah et al., 2003). In primates and rodents, aversive conditioning is defined as the reduction of prediction error (Rescorla & Wagner, 1972) associated with neural plasticity in the basolateral nuclei of the ventral amygdala, which receive sensory input from thalamic nuclei and sensory cerebral cortex, and project to the centromedial nuclei (Quirk et al., 1995; Fanselow &

LeDoux, 1999; Freese & Amaral, 2009; McHugh et al., 2014). The centromedial nuclei in the dorsal amygdala then project to the hypothalamus and brainstem nuclei and regulate observable physiological fear responses (LeDoux & Schiller, 2009). Extinction learning is conceptualized as learning of a novel association between external stimuli and the absence of threat, which inhibits behavioral responses triggered by previously learned associations (Bouton, 2004). In non-human primates and rodents, the circuit that mediates context-dependent inhibition has been shown to engage the hippocampus, which relays information to the centromedial amygdala via the ventromedial prefrontal cortex (Herry et al., 2008; Milad & Quirk, 2010; Tovote et al., 2015). While the circuits underlying fear conditioning and extinction have been delineated in animals, neuroimaging studies in humans have focused on specific regions of interest rather than on neural networks, functional connectivity, or network dynamics.

To date, studies on humans have shown that the amygdala, anterior cingulate cortex, and anterior insula are engaged during conditioning (LaBar et al., 1995; Büchel et al., 1998; Knight et al., 2003), that amygdala activation correlates with skin conductance response (SCR; LaBar et al., 1998), that the hippocampus is activated during context conditioning (Marschner et al., 2008), that ventromedial prefrontal cortex activity is related to extinction (Phelps et al., 2004; Milad et al., 2007), and that regions associated with extinction learning are de-activated during fear acquisition (Fullana et al., 2015). To date, however, human studies have not yet delineated the essential functional networks and their dynamics, i.e., they have not examined the dynamics of interregional functional interactions, which occur in the absence of significant changes in mean activity and which might prove crucial for identifying the subtle network changes underlying pathological fear learning (Grady et al., 1998; McIntosh et al., 1994).

The objective of this study was to delineate the functional networks, which subserve differential aversive delay conditioning and extinction in healthy humans, and to investigate the dynamic interactions between the delineated networks, using independent component analysis (ICA). ICA utilizes higher-order statistics to uncover the hidden sources (or independent components) that jointly contribute to a complex, measured signal, such as functional magnetic resonance imaging (fMRI). ICA results in an un-mixing of the contribution of different independent spatial components to the fMRI signal, which are interpreted as functional networks. Each component contains two anti-correlated patterns of brain activity, as well as the time course of their competition. The spatial-temporal structure of the components makes ICA an ideal method for investigation of the temporal dynamics of functional networks (Calhoun et al., 2009).

We used fMRI and a differential A-B-A-B conditioning and extinction paradigm that included repeated context-dependent reinforcement of conditioned stimuli (CS; partially reinforced: CS+, non-reinforced: CS-), in order to capture neural processes during initial and repeated phases of learning. By using a repeated learning paradigm, we were able to assess the effects of changes in threat contingencies as well as separate activity related to initial and sustained learning processes. Based on the aforementioned animal and human neuroimaging studies, we hypothesized (i) that an amygdala-based thalamo-cortical network would be engaged during conditioning, (ii) that a hippocampal-prefrontal network would be activated during extinction, and (iii) that these networks would dynamically interact with each other during fear acquisition and extinction. Given the recent finding that brain areas associated with extinction are de-activated during conditioning (Fullana et al., 2015), we expected to find evidence that the amygdala-thalamo-cortical network would be anti-correlated with the hippocampal-prefrontal network, either directly in a single component or in the temporal activation patterns.

Materials and Methods

Participants

30 right-handed adults (15 females, mean age = 26 years, age range = 21-34 years) with normal or corrected to normal vision took part in the experiment, which was approved by the Human Ethics Research Committee of the University of Queensland, after giving written consent. All participants were screened for neuropsychological disorders, brain damage, and substance abuse. Images were acquired with a Siemens Magnetom Trio 3T scanner and a 32-channel head coil at the Centre for Advanced Imaging, the University of Queensland.

Procedure

Participants took part in a partially reinforced, differential fear conditioning experiment, in which two visual stimuli (a triangle and a circle) were repeatedly presented in randomized order. Stimuli were presented in each of two contexts (blue or orange background) and contexts alternated between experimental blocks (A-B-A-B paradigm). Participants were asked to identify the stimuli by pressing one of two buttons with the second and third digit of their right hand. One of two conditioned stimuli (CS+) was paired with electro-dermal stimulation (unconditioned stimulus, UCS) in one of two contexts (danger context) but not the other (safe context), while the other stimulus (CS-) was never paired with stimulation. Stimuli and contexts were randomly assigned and assignments were counterbalanced across individuals. Each block started with 15s of background presentation to allow for the electrodermal response to settle and the participants to habituate. During each experimental block, following 1s of background, 20 stimuli (10 CS+, 10 CS-) were presented for 3s and followed by 14s of background in a randomized order. All stimuli were presented using Presentation software (Neurobehavioral Systems, Inc.) and projected onto a screen, which could be viewed with a mirror attached to the head coil.

Sixty percent of CS+ presentations co-terminated with a 50ms electro-dermal stimulation using two pre-gelled carbon snap electrodes attached to the right wrist (EL508, Biopac Systems, Inc.). Prior to scanning, stimulation strength was adjusted to individual tolerances following established procedures (LaBar et al., 1998) to ensure that stimulation was highly uncomfortable but not painful. Stimulation was administered using a STIMISOC isolator connected to a STM100C stimulator, which was attached to a MP150 (Biopac Systems, Inc.).

Skin conductance responses (SCRs) were sampled at 1kHz using pre-gelled carbon snap electrodes (EL508, Biopac Systems, Inc.) attached to the medial phalanges of the second and third digits of the left hand and connected to an EDA100C module attached to a MP150 (Biopac Systems, Inc.). SCRs were defined as the peak response of the low-pass filtered (0.1 Hz) electro-dermal activity occurring within 1-4s after the onset of the conditioned stimuli (Lockhardt, 1966). SCRs below 0.02 μS were excluded from the analysis.

Image Acquisition & Pre-processing

For each participant, a T1-weighted volumetric anatomical MRI was acquired with the following parameters: 176 slices sagittal acquisition MP2-RAGE; 1 mm^3 isotropic volume; repetition time (TR) = 4000 msec; echo time (TE) = 2.89 msec; flip angle = 6°; FOV = 256 mm, GRAPPA acceleration factor = 3. Functional images were acquired using a T2*-weighted echo-planar image pulse sequence with the following parameters: 45 slices; 2.7 mm slice thickness; voxel size = 2.5 x 2.5 x 2.7 mm; TR = 3000 msec; TE = 30 msec; FOV = 192 mm; flip angle = 90°. Brain activation was assessed using the blood oxygenation level dependent (BOLD) effect (Ogawa et al., 1990). For functional analysis, T2*-weighted images were pre-processed with Statistical Parametric Mapping software (SPM8; http://www.fil.ion.ucl.ac.uk/spm). Images were realigned to the mean image for head-motion correction and then spatially normalized into a standard stereotaxic space with voxel size of 2 mm^3 (Montreal Neurological Institute template) using segmented white and gray matter T1 maps. Head movement and rotation in the three dimensions did not exceed 1 mm and no dataset had to be excluded from analysis. Finally, the functional images were spatially smoothed with a 6-mm full width half maximum Gaussian kernel. All subsequent analysis of fMRI data is based on non-reinforced trials.

Independent Component Analysis

Following pre-processing, functional networks were identified with group independent component analysis (ICA) using the Group ICA of fMRI Toolbox (GIFT; http://mialab.mrn.org/software/gift/index.html). Individual images were first normalized to their mean intensity and then concatenated across time. The optimal number of independent components was estimated to be 32 using the minimum description length algorithm (Li et al., 2007). After data reduction with principal component analysis, 32 independent components (ICs) were identified using the infomax algorithm (Bell & Sejnowski, 1995). To estimate the stability of ICs, this analysis was repeated 20 times using ICASSO (Hirnberg et al., 2004). Only those ICs with a stability index larger than 0.95 were selected for further analysis. Finally, GICA back-reconstruction was applied to estimate the spatial maps and time courses of each IC for each participant using dual regression (Calhoun et al., 2001).

To identify task-relatedness of ICs, a general linear model (GLM) was fitted to each IC's time course. First, subject-specific regressors for each combination of stimulus and context were created for each of four imaging runs in SPM8 using convolution of a canonical hemodynamic response function with the stimulus onsets. Then, the beta-estimates of each regressor in the GLM that best predicted the back-reconstructed IC time course were estimated. Finally, a 2 x 2 x 2 ANOVA on beta-estimates with the factors stimulus (CS+, CS-), context (acquisition, extinction), and time (initial or repeated presentation) was used to identify significant differences in functional connectivity between CS+ and CS- presentations for each context. For display purposes only, the sign of negative task-related beta estimates was flipped and the related negative network was plotted as a positive network and vice versa. Temporal dynamics were assessed using calibrated back-reconstructed time courses. For each participant, the dwell time, i.e., the number of TRs, associated with one network (positive values) or the other (negative values) was calculated and averaged across participants for each experimental block.

Results

Electrophysiological evidence of successful differential fear conditioning was provided by a 2 x 2 x 2 analysis of variance of the SCRs with factors stimulus (CS+, CS-), context (acquisition, extinction), and time (early, late presentations) that yielded significant main effects for the factors stimulus ($F(1,1) = 5.4$, $p < 0.05$) and context ($F(1,1) = 17.2$, $p < 0.001$). Repeated two-sided t-tests demonstrated significant differences in participants' SCRs to CS+ and CS- presentations as well as to CS presentations during the acquisition and extinction phases (all t(28) > 2.1, $p < 0.05$; see Figure 1).

(INSERT FIGURE 1 HERE)

ICA yielded 32 independent components (ICs), out of which two stable ICs included the hypothesized regions and showed a significant interaction between cue and context. Each IC included two anti-correlated networks arbitrarily differentiated by their sign, i.e., when the positive network is activated, the negative network is deactivated and vice versa. Each IC's correlation with the task was assessed by fitting a general linear model with task-related regressors that have been convolved with a standard hemodynamic response function to the IC's time course.

(INSERT FIGURE 2 HERE)

The first component (IC5) showed a significant three-way interaction between cue, context, and time (F(1) = 4.81, $p = 0.03$). IC5's positive network was positively correlated with the CS+ during the initial, but not repeated acquisition, and included the insula, dorsal amygdala, thalamus, brainstem, and anterior hippocampus. This network reflects the initial acquisition of a differential expectation of aversive reinforcement and can therefore be considered a rapid fear-learning network (see Figure 2).

(INSERT FIGURE 3 HERE)

The second component (IC15) showed a significant interaction between cue and context and was significantly more related to CS+ than CS- processing during initial and repeated fear acquisition, but not during extinction (F(1) = 6.46, p = 0.01). The positive network of IC15 included regions previously shown to be involved in fear learning, such as the ventral amygdala, anterior hippocampus, temporal pole, as well as middle frontal and inferior parietal cortex. The negative network of IC15 included areas that were negatively correlated with fear acquisition and that have previously been shown to be activated during extinction-learning, such as the ventral striatum, posterior hippocampus, ventromedial prefrontal cortex, anterior cingulate cortex, frontal operculum, and posterior cingulate cortex (Phelps et al., 2004; see Figure 3). Together, IC15's networks provide evidence for the view that fear and extinction learning engage two separate but anti-correlated networks. To analyze the temporal dynamics of these networks, we calculated their individual dwell time, i.e., the time each individual's brain spent in one of the two anti-correlated networks. Post-hoc analysis of IC15's time courses showed context-dependent changes in dwell time. On average, the results show a non-significant increase in dwell time by 0.44% in the fear acquisition network between acquisition phases, as well as a non-significant decrease in dwell time by 0.34% in the fear acquisition network between extinction phases (all $t_{29}<2$; see Figure 4). These results suggest that changes in associative contingencies affect the dwell time and hence bias the dynamic competition between networks.

(INSERT FIGURE 4 HERE)

Both components (IC5 and IC15) included connectivity with the amygdala during fear acquisition. Previous studies found differences in activation of the dorsal (including the superficial nuclei) and the ventral regions (including the basolateral nuclei) related to initial and sustained fear acquisition, respectively (Morris et al., 2001). A post-hoc comparison of the networks revealed that IC5 engaged the dorsal region of the amygdala, whereas IC15 engaged the ventral region of the amygdala (see Figure 5). In other words, while both amygdala regions were related to fear acquisition, the dorsal region was only engaged during the initial acquisition, whereas the ventral region was engaged during the initial and the repeated acquisition. These results replicate previous findings about local amygdala activations during initial and sustained conditioning and confirm the validity of the results (Morris et al., 2001).

(INSERT FIGURE 5 HERE)

Discussion

The results of this study demonstrate that two learning-related brain networks dynamically compete with each other during associative learning and that the outcome of this competition distinguishes conditioning from extinction. The results specifically show that in the human brain, activity in the amygdala-thalamo-cortical network associated with aversive learning is anti-correlated with activity in the hippocampal-prefrontal network associated with extinction learning. Our results therefore replicate and extend previous findings about the activation of brain regions during fear conditioning and extinction (Büchel et al., 1998; Marschner et al., 1998; Phelps et al., 2004; for a meta-analysis of neuroimaging studies, see Fullana et al., 2015). Furthermore, the analysis of the temporal dynamics revealed that the oscillation between these two network states is sensitive to changes in associative contingencies, such that the net outcome of their competition predicts the difference between fear conditioning and extinction. In other words, whether a cue-context combination is being associated with an aversive outcome or not seems to depend on the relative time the brain spends in one network state over the other.

The evidence of dynamic oscillations between learning-related neural networks lends support to the view that the human brain forms a meta-stable system, in which transient networks compete with each other (Shanahan, 2010; Deco & Jirsa, 2012; Mazzucato et al., 2015). The dynamic view of brain connectivity aligns with the proposal that competition is the underlying brain mechanism by which neural resources are allocated to different learning systems without prior knowledge about the nature of the learning problem (Fanselow, 2010). Our findings add to the evidence that learning systems compete with each other and further suggest that not only is there competition between learning systems but that the competition between learning systems is sensitive to changes in associative contingencies. Previous research suggests that the transient networks in a meta-stable system are stabilized by sensory input (Churchland et al., 2010; Litwin-Kumar & Doiron, 2012; Ponce-Alvarez et al., 2015). Our results show that changes in associative contingencies bias the competition towards a particular state resulting in an increased net dwell time in the respective network. In other words, a cue-context combination that is presented with an unconditioned stimulus biases the competition between meta-stable networks and leads to an increased net dwell time in the amygdala-thalamo-cortical network whereas the absence of an unconditioned stimulus leads to an decreased net dwell time in the amygdala-thalamo-cortical network.

Interestingly, our results show an overall longer net dwell time for the amygdala-thalamo-cortical network compared to the hippocampal-prefrontal network. This finding suggests that the competition between the conditioning and the extinction networks might initially be biased towards conditioning. Such a competition bias might possibly reflect the result of evolutionary pressure to minimize losses due to the higher prize for error in dangerous rather than safe situations. This interpretation is consistent with the idea that the human brain is constantly optimizing its organization towards reducing surprise (Grossberg, 2009; Friston, 2010; Clark, 2013), and thus effectively forming a survival optimization system (Mobbs et al., 2015).

Our findings demonstrate the dynamic competition between learning-related networks in healthy young adults and suggest that flexible modulation of network dynamics is *essential* for adaptive behaviour. The implications of these findings extend to clinical conditions characterized by excessive or chronic fear. Many learning-related disorders, such as PTSD, can be characterized by a difficulty to engage a particular type of learning, such as extinction learning (Kim et al., 2011; Jovanovic et al., 2012). In this context, our findings suggest that maladaptive associative learning might be the result of dysfunctional network competition. As such, characterizations of maladaptive dynamics of the conditioning and extinction networks may be essential to shed light on learning-related pathogenesis and guide the development of clinical biomarkers of learning-related disorders, such as PTSD (Uhlhaas & Singer, 2012; Michopoulos et al., 2015; Kringelbach et al., 2015).

Acknowledgements:

We thank Aiman Al Najjar and Nicole Atcheson for assistance with data collection. This work was funded by the Australian Research Council Special Research Initiative: Science of Learning Research Centre (project number SR120300015).

References

Bell AJ, Sejnowski TJ. 1995. An information-maximization approach to blind separation and blind deconvolution. Neural Comput. 7:1129–59.

Bouton ME. 2004. Context and behavioral processes in extinction. Learn Mem. 11: 485-94.

Büchel C, Morris J, Dolan RJ, Friston KJ. 1998. Brain systems mediating aversive conditioning: an event-related fMRI study. Neuron. 20: 947-57.

Calhoun VD, Adalı T, Pearlson GD, Pekar JJ. 2001. A method for making group inferences from functional MRI data using independent component analysis. Hum Brain Mapp. 14: 140–51.

Calhoun VD, Liu J, Adalı T. 2009. A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. Neuroimage. 45(1 Suppl): S163-72.

Calhoun VD, Miller R, Pearlson G, Adalı T. 2014. The chronnectome: time-varying connectivity networks as the next frontier in fMRI data discovery. Neuron. 84: 262-74.

Churchland MM, Yu BM, Cunningham JP, Sugrue LP, Cohen MR, Corrado GS, Newsome WT, Clark AM, Hosseini P, Scott BB, Bradley DC, Smith MA, Kohn A, Movshon JA, Armstrong KM, Moore T, Chang SW, Snyder LH, Lisberger SG, Priebe NJ, Finn IM, Ferster D, Ryu SI, Santhanam G, Sahani M, Shenoy KV. 2010. Stimulus onset quenches neural variability: a widespread cortical phenomenon. Nat Neurosci. 13(3): 369-78.

Clark A. 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav Brain Sci. 36(3):181-204.

Deco G, Jirsa VK. 2012. Ongoing cortical activity at rest: criticality, multistability, and ghost attractors. J Neurosci. 32(10): 3366-75.

Fanselow MS. 1994. Neural organization of the defensive behaviour system responsible for fear. Psychonomic Bulletin and Review. 1: 429-38.

Fanselow MS. 2010. From contextual fear to a dynamic view of memory systems. Trends Cogn Sci. 14(1): 7-15.

Fanselow MS, LeDoux JE. 1999. Why we think plasticity underlying Pavlovian fear conditioning occurs in the basolateral amygdala. Neuron. 23: 229-32.

Freese J, Amaral D. 2009. Neuroanatomy of the Primate Amygdala. In: Whalen PE & Phelps E, editors. The human amygdala. New York: Guilford. p 3–42.

Friston K. 2010. The free-energy principle: a unified brain theory? Nat Rev Neurosci. 11(2):127-38.

Fullana MA, Harrison BJ, Soriano-Mas C, Vervliet B, Cardoner N, Àvila-Parcet A, Radua J. 2015. Neural signatures of human fear conditioning: an updated and extended meta-analysis of fMRI studies. Mol Psychiatry. doi: 10.1038/mp.2015.88.

Grady CL, McIntosh AR, Bookstein F, Horwitz B, Rapoport SI, Haxby JV. 1998. Age-related changes in regional cerebral blood flow during working memory for faces. Neuroimage. 8(4): 409-25.

Grossberg S. 2009. Cortical and subcortical predictive dynamics and learning during perception, cognition, emotion and action. Philos Trans R Soc Lond B Biol Sci. 364(1521): 1223-34.

Herry C, Ciocchi S, Senn V, Demmou L, Müller C, Lüthi M. 2008. Switching on and off fear by distinct neuronal circuits. Nature. 454: 600-6.

Himberg J, Hyvarinen A, Esposito F. 2004. Validating the independent components of neuroimaging time series via clustering and visualization. Neuroimage. 22: 1214–22.

Jovanovic T, Kazama A, Bachevalier J, Davis M. 2012. Impaired safety signal learning may be a biomarker of PTSD. Neuropharmacology. 62(2): 695-704.

Kim MJ, Loucks RA, Palmer AL, Brown AC, Solomon KM, Marchante AN, Whalen PJ. 2011. The structural and functional connectivity of the amygdala: from normal emotion to pathological anxiety. Behav Brain Res. 223: 403-20.

Knight DC, Smith CN, Cheng DT, Stein EA, Helmstetter FJ. 2004. Amygdala and hippocampal activity during acquisition and extinction of human fear conditioning. Cognitive, Affective, & Behavioral Neuroscience. 4(3): 317-25.

Kringelbach ML, McIntosh AR, Ritter P, Jirsa VK, Deco G. 2015. The Rediscovery of Slowness: Exploring the Timing of Cognition. Trends Cogn Sci. 19: 616-28.

LaBar KS, LeDoux JE, Spencer DD, Phelps EA. 1995. Impaired fear conditioning following unilateral temporal lobectomy in humans. J Neurosci. 15(10): 6846-55.

LaBar KS, Gatenby JC, Gore JC, LeDoux JE, Phelps EA. 1998. Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fMRI study. Neuron. 20: 937-45.

Lang PJ, Davis M, Öhman A. 2001. Fear and anxiety: Animal models and human cognitive psychophysiology. Journal of Affective Disorders. 61: 137-59.

LeDoux JE. 2000. Emotion circuits in the brain. Annual Review of Neuroscience. 23: 155-84.

LeDoux JE, Schiller. 2009. The human amygdala: Insights from other animals. In: Whalen PE & Phelps E, editors. The human amygdala. New York: Guilford. p 43-60.

Li YO, Adali T, Calhoun VD. 2007. Estimating the number of independent components for functional magnetic resonance imaging data. Hum Brain Mapp. 28: 1251–66.

Li X, Zhu D, Jiang X, Jin C, Zhang X, Guo L, Zhang J, Hu X, Li L, Liu T. 2014. Dynamic functional connectomics signatures for characterization and differentiation of PTSD patients. Hum Brain Mapp. 35(4): 1761-78.

Litwin-Kumar A, Doiron B. 2012. Slow dynamics and high variability in balanced cortical networks with clustered connections. Nat Neurosci. 15(11): 1498-505.

Lockhardt RA. 1966. Comments regarding multiple response phenomena in long interstimulus interval conditioning. Psychophys. 3: 108-14.

Marschner A, Kalisch R, Vervliet B, Vansteenwegen D, Büchel C. 2008. Dissociable roles for the hippocampus and the amygdala in human cued versus context fear conditioning. J Neurosci. 28: 9030-6.

Mazzucato L, Fontanini A, La Camera G. 2015. Dynamics of multistable states during ongoing and evoked cortical activity. J Neurosci. 35(21): 8214-31.

McHugh SB, Barkus C, Huber A, Capitão L, Lima J, Lowry JP, Bannerman DM. 2014. Aversive prediction error signals in the amygdala. J Neurosci. 34: 9024-33.

McIntosh AR, Grady CL, Ungerleider LG, Haxby JV, Rapoport SI, Horwitz B. 1994. Network analysis of cortical visual pathways mapped with PET. J Neurosci. 14(2): 655-66.

Michopoulos V, Norrholm SD, Jovanovic T. 2015. Diagnostic Biomarkers for Posttraumatic Stress Disorder: Promising Horizons from Translational Neuroscience Research. Biol Psychiatry. 78(5): 344-53.

Milad MR, Wright CI, Orr SP, Pitman RK, Quirk GJ, Rauch SL (2007). Recall of fear extinction in humans activates the ventromedial prefrontal cortex and hippocampus in concert. Biological Psychiatry. 62(5): 446-54.

Milad MR, Quirk GJ. 2010. Neurons in medial prefrontal cortex signal memory for fear extinction. Nature. 420, 70-4.

Mobbs D, Hagan CC, Dalgleish T, Silston B, Prévost C. 2015. The ecology of human fear: survival optimization and the nervous system. Front Neurosci. 9: 55.

Morris JS, Büchel C, Dolan RJ. 2001. Parallel neural responses in amygdala subregions and sensory cortex during implicit fear conditioning. Neuroimage. 13: 1044-52.

Ogawa S, Lee TM, Kay AR, Tank DW. 1990. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. Proc Natl Acad Sci USA. 87: 9868–72.

Panzeri S, Macke JH, Gross J, Kayser C. 2015. Neural population coding: combining insights from microscopic and mass signals. Trends Cogn Sci. 19(3): 162–72.

Phelps EA, Delgado MR, Nearing KI, LeDoux JE. 2004. Extinction learning in humans: role of the amygdala and vmPFC. Neuron. 43: 897-905.

Ponce-Alvarez A, He BJ, Hagmann P, Deco G. 2015. Task-Driven Activity Reduces the Cortical Activity Space of the Brain: Experiment and Whole-Brain Modeling. PLoS Comput Biol. 11(8): e1004445.

Quirk GJ, Repa C, LeDoux JE. 1995. Fear conditioning enhances short-latency auditory responses of lateral amygdala neurons: parallel recordings in the freely behaving rat. Neuron. 15: 1029–39.

Rescorla RA, Wagner AR. 1972. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: Black AH & Prokasy WF, editors. Classical Conditioning II. Appleton-Century-Crofts. p 64-99.

Sah P, Faber ES, Lopez De Armentia M, Power J. 2003. The amygdaloid complex: anatomy and physiology. Physiol Rev. 83, 803-34.

Shanahan M. 2010. Metastable chimera states in community-structured oscillator networks. Chaos. 20(1): 013108.

Tovote P, Fadok JP, Lüthi A. 2015. Neuronal circuits for fear and anxiety. Nat Rev Neurosci. 16: 317-31.

Uhlhaas PJ, Singer W. 2012. Neuronal Dynamics and Neuropsychiatric Disorders: Toward a Translational Paradigm for Dysfunctional Large-Scale Networks. Neuron. 75: 963-80.

Zalesky A, Fornito A, Cocchi L, Gollo LL, Breakspear M. 2014. Time-resolved resting-state brain networks. Proc Natl Acad Sci USA. 111: 10341-6

Figure Legends:

Figure 1: Skin conductance responses (SCRs). Bar graphs show group mean SCRs in response to presentations of CS+ (left) and CS- (right) during acquisition / ACQ (dark grey) and extinction / EXT learning (light grey) for early phases, i.e., initial acquisition and extinction, (top) and late phases, i.e., repeated acquisition and extinction (bottom). Stars indicate significant main effects of stimulus and context.

Figure 2: Rapid fear-learning network (IC5). Left: Insula, amygdala, ventromedial prefrontal cortex, and hippocampus form a network (warm colors) that is anti-correlated with a network that includes orbitofrontal cortex, striatum, brainstem, and cerebellum (cool colors). Right: The graph plots group means and SEMs of task-relatedness (beta estimates) for CS+ (gray) and CS- (white) during initial and repeated fear acquisition (ACQ/RACQ) and extinction (EXT/REXT). ANOVA shows that the initial fear-learning network is significantly more related to the CS+ during the early (but not the later) stages of fear acquisition than the CS-.

Figure 3: Sustained fear-learning network (IC15). Left: Amygdala-thalamo-cortical network (warm colors) that is anti-correlated with a hippocampal-prefrontal network (cool colors). Right: The graph plots group means and SEMs of task-relatedness (beta estimates) for CS+ (gray) and CS- (white) during initial and repeated fear acquisition (ACQ) and extinction (EXT; beta estimates are collapsed across initial and repeated phases). ANOVA shows that the amygdala-thalamo-cortical fear conditioning network is significantly more related to the CS+ than the CS- during fear acquisition but not extinction.

Figure 4: Left: Dynamic competition between networks. The line plots of time courses of IC15 (group ICA) show temporal oscillations between the two anti-correlated sub-networks of for acquisition (top row) and extinction training (bottom row). Positive values show dwell time in the amygdala-thalamo-cortical fear conditioning network, whereas negative values show dwell time in the hippocampal-prefrontal extinction network. Dotted lines show oscillations for initial phases of training, whereas solid lines show oscillations for repeated phases of training in the Acquisition-Extinction-Acquisition-Extinction paradigm. Right: The bar graphs show the mean differences and SEMs in network dwell times between the amygdala-thalamo-cortical and the hippocampal-prefrontal networks across all participants (back-reconstructed ICs), which increase from initial to repeated acquisition and decrease from initial to repeated extinction.

Figure 5: Dissociation between ventral and dorsal amygdala connectivity. Bottom: Coronal (y = -3) and sagittal (x = 23, MNI coordinates) slices show peak amygdala activations within the positive networks of IC5 (warm colors) and IC15 (cool colors). Top: Magnified sections show that IC5 (warm colors) engages dorsal amygdala during initial fear acquisition whereas IC15 (cool colors) engages ventral amygdala during initial and repeated fear acquisition.

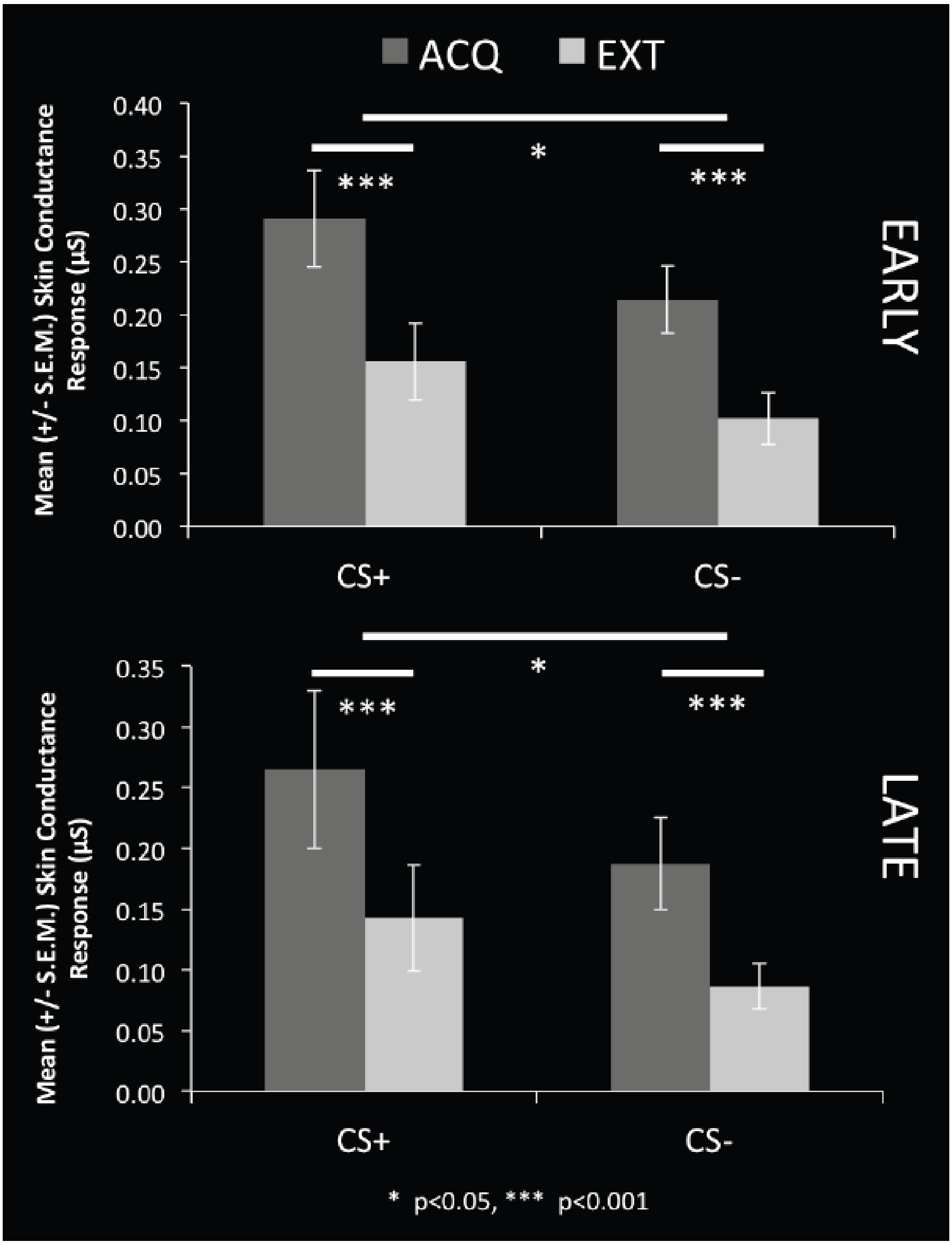
ACQ
EXT
Mean (+/- S.E.M.) Skin Conductance Response (µS)
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
*
***
***
EARLY
CS+
CS-
Mean (+/- S.E.M.) Skin Conductance Response (µS)
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
*
***
***
LATE
CS+
CS-
* p<0.05, *** p<0.001

Figure 1

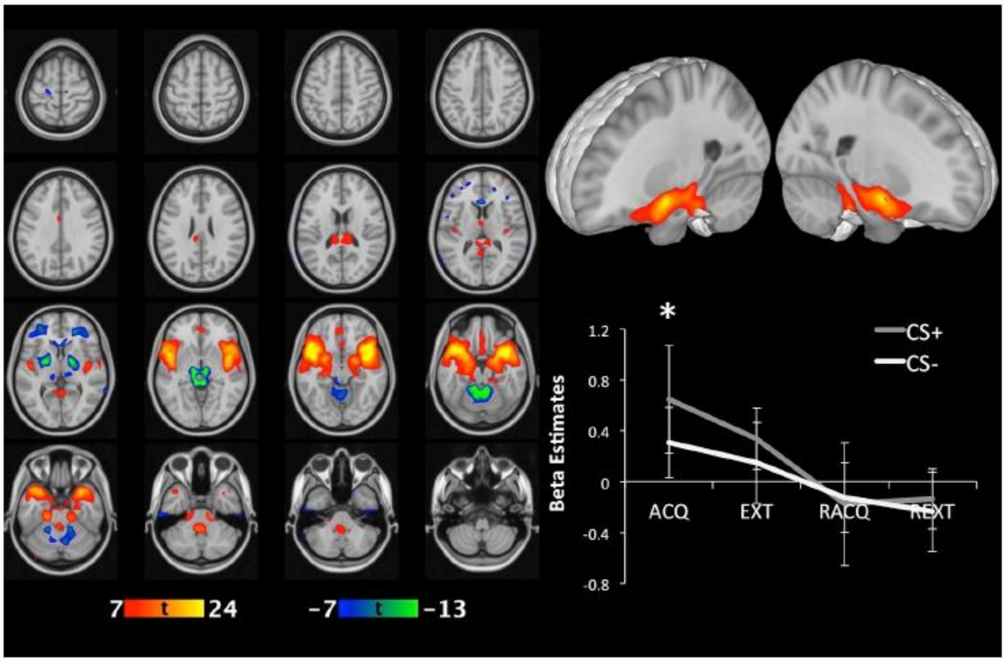
7 t 24
-7 t -13
*
1.2
0.8
0.4
0
-0.4
-0.8
Beta Estimates
ACQ
EXT
RACQ
REXT
CS+
CS-

Figure 2

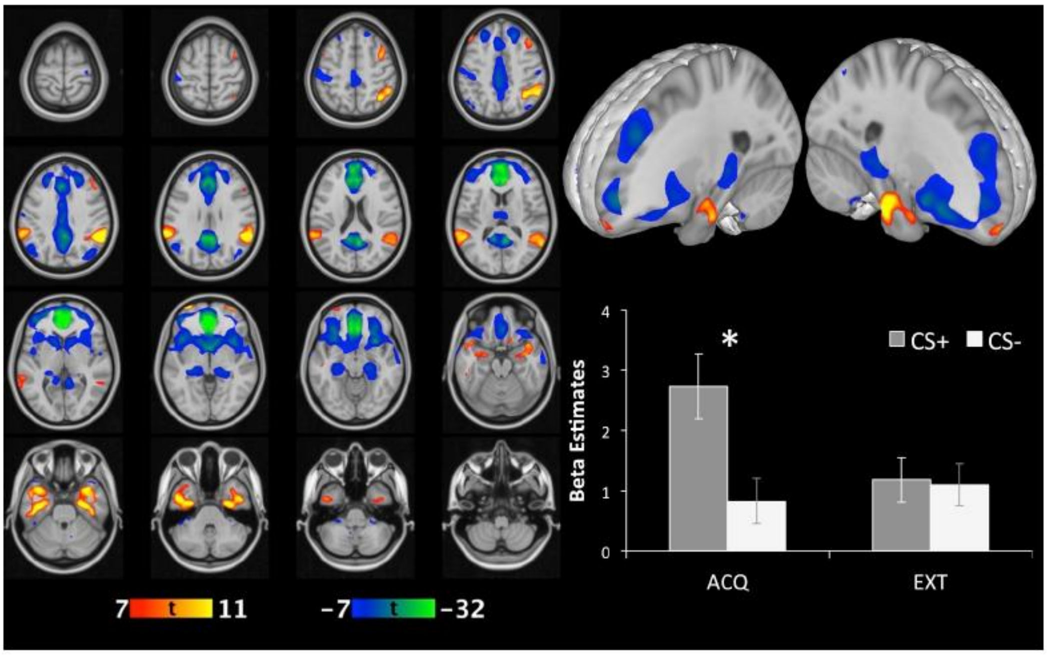

Figure 3

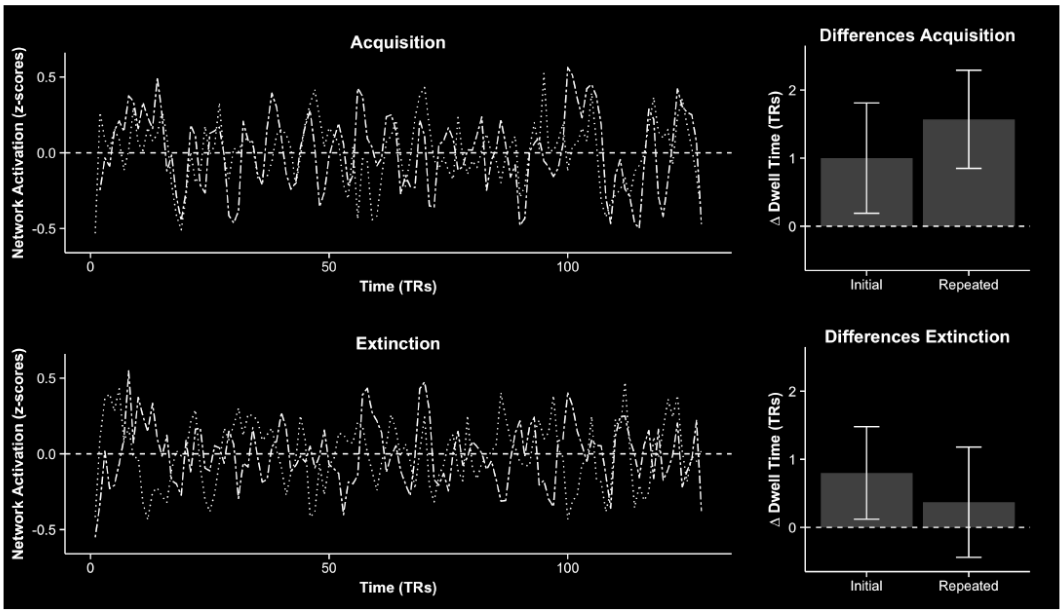
Acquisition
Network Activation (z-scores)
0.5
0.0
-0.5
0
50
100
Time (TRs)
Differences Acquisition
Δ Dwell Time (TRs)
2
1
0
Initial
Repeated
Extinction
Network Activation (z-scores)
0.5
0.0
-0.5
0
50
100
Time (TRs)
Differences Extinction
Δ Dwell Time (TRs)
2
1
0
Initial
Repeated

Figure 4

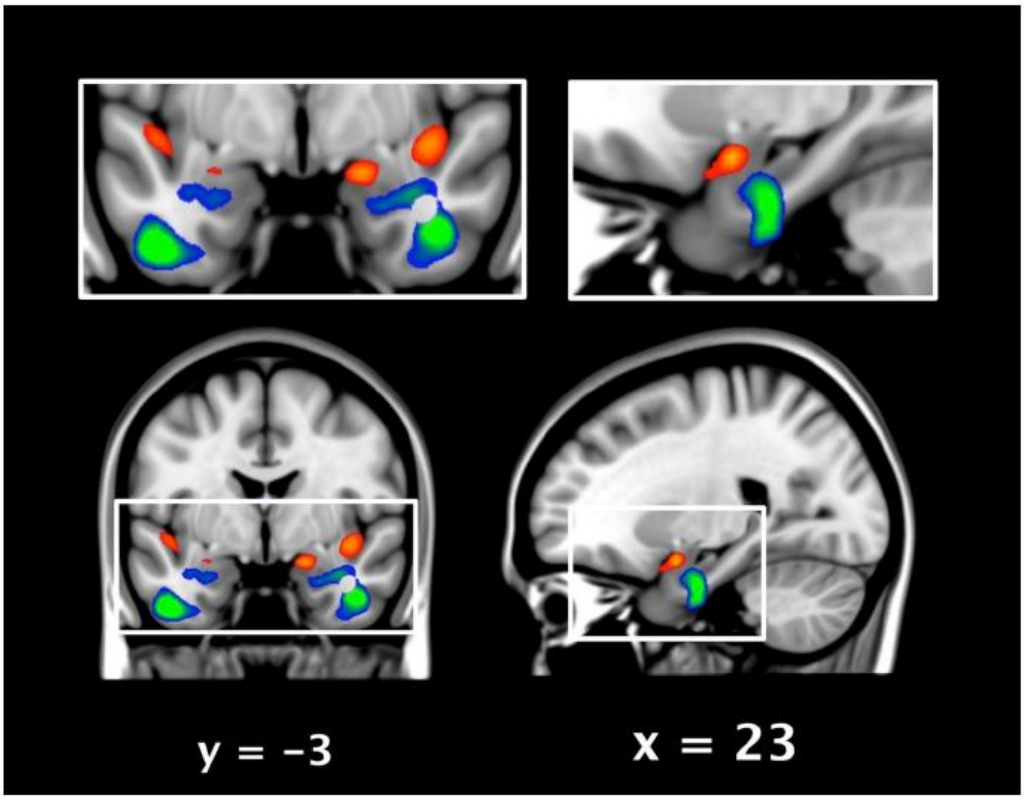

Figure 5

Highlights

- Networks underlying fear acquisition and extinction are anti-correlated
- Competition between amygdala-thalamo-cortical and hippocampal-prefrontal networks
- Meta-stable dynamic learning is stabilized by external associative contingencies
- Flexible modulation of network dynamics is essential for adaptive behaviour
- Maladaptive associative learning might be result of dysfunctional network dynamics