

Classification:

Major-Biological Sciences

Minor-Psychological and Cognitive Sciences

Regional inactivations of primate ventral prefrontal cortex reveal two distinct mechanisms underlying negative bias in decision making

Hannah F. Clarke^{1,2*}, Nicole K. Horst^{1,3}, Angela C. Roberts^{1,2}.

1. Department of Physiology, Development and Neuroscience, University of Cambridge, Downing Street, Cambridge, Cambridgeshire, CB23 3DY, UK

2. Behavioural and Clinical Neurosciences Institute, University of Cambridge, Downing Street, Cambridge, Cambridgeshire, CB23 3EB, UK

3. Department of Psychology, University of Cambridge, Downing Street, Cambridge, Cambridgeshire, CB23 3EB, UK

* corresponding author:

Email: hfc23@cam.ac.uk

Short title: Prefrontal fractionation of anxiety

Key words: anxiety, orbitofrontal cortex, ventrolateral prefrontal cortex, negative bias

Significance Statement

Fear of negative outcomes has a powerful adverse influence on decision making in anxiety disorders. While neuroimaging studies of patients with anxiety disorders have revealed dysregulation in numerous frontal brain regions including the orbitofrontal and ventrolateral prefrontal cortex, the causal involvement of this dysregulation is unknown. Here we demonstrate that in the marmoset monkey, inactivation of anterior orbitofrontal or ventrolateral prefrontal cortex increases negative bias in decision making via two distinct cognitive mechanisms – elevated uncertainty and attentional disruption, respectively. These findings provide the first direct evidence that dysregulation of distinct neurocognitive mechanisms within the prefrontal cortex may underlie the mixed aetiology of anxiety disorders. Such insight will allow the development of more precise diagnostics and individually tailored therapeutic approaches.

Abstract

Dysregulation of the orbitofrontal (OFC) and ventrolateral (vlPFC) prefrontal cortices is implicated in anxiety and mood disorders but the specific contributions of each region are unknown, including how they gate the impact of threat on decision making. To address this, the effects of GABAergic inactivation of these regions were studied in marmoset monkeys performing an instrumental approach-avoidance decision-making task that is sensitive to changes in anxiety. Inactivation of either region induced a negative bias away from punishment that could be ameliorated with anxiolytic treatment. However, whereas the effects of vlPFC inactivation on punishment avoidance were seen immediately, those of OFC inactivation were delayed and their expression dependent upon an amygdala-anterior hippocampal circuit. We propose that these negative biases result from deficits in attentional control and punishment prediction, respectively and that they provide the basis for understanding how distinct regional prefrontal dysregulation contributes to the heterogeneity of anxiety disorders with implications for cognitive-behavioural treatment strategies.

\body

Introduction

Sensitivity to threat, and the appropriate interpretation of potential threat, is crucial for an organism to survive and make optimal decisions with respect to its environment. Overestimation of threat, and hypersensitivity to negative emotional information are known to inappropriately impact cost/benefit decision making in patients suffering from anxiety and depression (1, 2). This hypersensitivity is thought to be due to dysregulation within the

prefrontal cortex (PFC), but how the PFC contributes to aversive processing and how it gates the impact of negative emotional information on decision making is still poorly understood.

There are a number of distinct brain regions within the PFC that are dysregulated in anxiety and mood disorders including the orbitofrontal (OFC), ventrolateral prefrontal (vlPFC) and medial prefrontal (mPFC) cortices (3–5). Of these, a region within mPFC (pregenual cingulate cortex) has been implicated in regulating negative emotional valence in decision making (6), but the contribution of the other regions remain unknown. Given the lifetime prevalence and economic cost of anxiety and depression (7), understanding how these distinct prefrontal subregions modulate the impact of emotion on decision making is crucial in order to identify how different types of prefrontal dysregulation contribute to the heterogeneity of anxiety and mood disorders and thus guide the development of personalised treatments. Despite the uncertainty regarding the rodent correlates of these other prefrontal regions, in particular vlPFC, there have been few studies investigating the selective contribution of these other prefrontal regions to negative decision making in primates, as most primate studies focus on reward guided decision making (8, 9) but see (10–13). However, we showed previously that selective excitotoxic lesions of either anterior OFC (antOFC; area 11) or vlPFC (area 12) heighten anxiety and pavlovian fear responses in marmoset monkeys, demonstrating that both regions contribute independently to the regulation of negative emotion (14), but their differential contribution and their involvement in modulating the impact of anxiety on decision making remains unknown. To address this, we developed an approach-avoidance conflict task suitable for marmoset monkeys, and used anatomically specific intracerebral infusions and anxiolytic drug treatment to determine how temporary inactivation of these regions affected cost/benefit decision making.

Results

Marmosets were trained to respond to two identical visual stimuli presented on each side of a touchscreen to gain access to reward (5s banana juice) that were on independent but identical variable-interval schedules (VI, 15s; Figure 1a). Thus the optimal strategy for maximising reward delivery is to respond relatively equally to both stimuli. Marmosets did this whilst showing a slight preference for one side over the other. See Figure 1b. Approximately once a week, responses on one of the stimuli also produced punishment, an aversive loud noise (0.3s, 117dB) on a leaner independent variable-interval schedule (VI 40s) that was superimposed upon the existing reward schedule. This punishment was always introduced onto the marmosets' 'preferred' side to avoid any spatial bias contributing to a punishment-induced bias. In the absence of reward, this loud noise, when paired with a neutral cue, has been shown to induce Pavlovian conditioned cardiovascular arousal and behavioral vigilance (14), indicative of its aversive properties. By combining the aversive noise with reward, this task measures the extent of avoidance of the punished schedule when this conflicts with a competing approach response for reward. In the absence of any brain manipulation or following saline infusions, the addition of the punishment on one side did not alter the animals' responding i.e they did not alter their behaviour to avoid it ($F_{1,7} = 0.16, P = 0.698$; see Figure 1b). Thus, it would appear that in the control condition, the animals still find the reward 'worth' responding for, despite the occasional punishment it incurs (14). Following implantation of intracerebral cannulae targeting either the antOFC (area 11) or vlPFC (area 12; Fig 1c,d), these cortical regions were then inactivated bilaterally with a GABA agonist mixture (0.5 μ l of 0.1 mM muscimol/1.0 mM baclofen) (15, 16) or saline, 20 min prior to reward-only or reward and punishment test sessions to determine their contribution to the integration of reward and punishment in decision making.

Inactivation of the vlPFC and antOFC induce a negative decision bias on different timescales

Inactivation of either the vIPFC or antOFC had no effect on responding when stimuli on both sides of the screen produced only reward: monkeys maintained the same level of responding and continued to respond to both sides equally. In contrast, when punishment was introduced on one side only, inactivation of either area induced a bias in responding away from the punished side. (Response biases were calculated as the ratio between the number of responses to the monkey's non-preferred side and the preferred side on an inactivation day and compared to the same measure on a saline infusion day). However, these effects differed as to whether they were seen on the infusion day (day 1/infusion day), or on the next day (day 2; three way interaction on sqrt transformed data on extent of bias, feedback(just reward, or reward and punishment on a given side) x day(day1, day 2) x group (antOFC, vIPFC), $F_{1,6} = 22.02$, $P = 0.003$; day x group (reward only), $F_{1,6} > 1$; day x group (punishment and reward combined), $F_{1,6} = 24.29$, $P = 0.003$; Fig 2a). Following control infusions of saline, all animals continued to respond for reward in the presence of punishment, whereas inactivation of the vIPFC resulted in a marked bias in responding away from the punished side and an increase in responding to the non-punished (reward only) side. This bias developed during the punished session (day 1) but no delayed/long lasting effects were seen the following day when punishment was absent (posthoc analysis: high bias day 1 vs no bias day 2, $t_3 = -3.996$, $P = 0.028$). In contrast, and similar to control infusions, inactivation of the antOFC had no effect on responding on the punished session. However, unlike control infusions or vIPFC inactivation, it did result in a profound bias away from the previously punished side on the following (reward-only) day (no bias day 1 vs high bias day 2, $t_3 = 3.264$, $P = 0.047$). Neither manipulation resulted in alterations in overall numbers of responses, ruling out any effects due to changes in Pavlovian driven punishment-induced suppression of responding (Fig 2b). The differential pattern of effects suggests that inactivation of the vIPFC altered the cost-benefit analysis at the time of the decision, while inactivation of the antOFC affected the

consolidation of a memory for the punishment that was sufficient to drive the anti-punishment bias the following day.

An anxiolytic abolished the antOFC and vlPFC inactivation-induced negative decision bias.

Given that permanent lesions of both these prefrontal regions enhance anxiety (14), it was possible that these inactivation-induced biases were due to an increase in anxiety-related behaviour. We therefore determined whether these biases away from punishment could be blocked by a known anxiolytic drug. Administration of diazepam (0.25mg/kg i.m. vs saline, Fig 3), 30 mins prior to the punished session (day 1) abolished not only the anti-punishment bias induced by vlPFC inactivation on the same day ($t_3 = 10.25$, $P = 0.002$) but also abolished the delayed (day 2) anti-punishment bias induced by antOFC inactivation ($t_3 = 10.404$, $P = 0.002$). Furthermore, diazepam also reduced the expression of the effects of day 1, antOFC inactivation when given before the reward-only, non-infusion session on day 2 ($t_3 = 10.04$, $P = 0.002$; see table S1 for infusion order). Together these findings suggest an anxiety-like state underpinned the effects of vlPFC inactivation on ‘on-line’ cost-benefit analysis and the effects of antOFC inactivation on a punishment memory.

Inactivation of the antOFC modulates a punishment memory within an amygdala-hippocampal circuit

The anti-punishment bias seen the day after antOFC inactivation suggests that learning to predict the punishment and the formation of a punishment memory is normally modulated by the antOFC. Two likely subcortical candidates for the location or expression of such a punishment memory are the amygdala and the anterior hippocampus. Both structures are

implicated in emotional processing and are connected to the antOFC and to each other (17). To investigate their contribution to the expression of the anti-punishment bias, we cannulated the amygdala (Fig 4a,c) and anterior hippocampus (Fig 4a,d) bilaterally in three of the antOFC-cannulated animals. AntOFC inactivation on the punishment day was followed the next day by inactivation of the amygdala bilaterally, hippocampus bilaterally, or both, unilaterally, in a crossed disconnection, before testing. These infusions were compared to the effects of identical amygdala/hippocampal manipulations the day after punishment, but in the absence of antOFC inactivation. Throughout these infusions, the bias away from the punished side induced by antOFC inactivation alone was still present and no different from its first presentation ($t_2 = 0.177$, $P = 0.876$). However, bilateral inactivation of either the anterior hippocampus ($t_2 = 11.808$, $P = 0.007$) or the amygdala ($t_2 = 10.185$, $P = 0.01$) on day 2, abolished the bias away from the punished side that was induced by antOFC inactivation on day 1, confirming the involvement of these structures in the expression of increased punishment sensitivity after antOFC inactivation. Disconnecting the amygdala from the hippocampus by inactivating one side of each structure in opposite hemispheres, whilst leaving the other sides intact, had the same effect as bilateral inactivations of either structure ($t_2 = 10.197$, $P = 0.009$), indicating that the expression of this behaviour is subserved by a functional amygdala-hippocampal circuit (Fig 4b).

Discussion

These results reveal the critical but dissociable roles played by the vIPFC and the antOFC in modulating the impact of threat-induced anxiety on instrumental cost-benefit decisions, and provide insight into the distinct contributions that each region makes to decision making. Inactivation of the vIPFC but not the antOFC increased punishment avoidance when making decisions between primary, unconditioned rewards and punishments. In contrast, antOFC

inactivation had no effect on the decision making process per se but affected the memory for punishment, increasing avoidance of the previously punished side the next day, in the absence of explicit punishment.

There have only been a handful of studies investigating the effects of neural interventions restricted to the vIPFC e.g. (18, 19) but we have previously demonstrated its importance for orienting/shifting attention to reward-relevant stimuli in marmosets (20), a finding supported by neuroimaging and electrophysiological studies implicating this region in attentional control (21, 22). We now propose that a deficit in attentional control underlies the increase in punishment avoidance induced by vIPFC inactivation. Specifically, we suggest that to make a cost-benefit decision, attention needs to be shifted away from the highly salient, negative outcome (the punishment), and towards the positive (rewarding) outcome, to facilitate comparison of the relative values of the reward and punishment. In the intact animal, the vIPFC provides top-down attentional control that allows the rewarding outcome to gain such attention, providing the animal with the opportunity to consider both reward and punishment when responding on that side. However, when an animal with an inactivated vIPFC responds to the punished side, the highly salient, aversive outcome captures their attention as before, but they fail to shift their attention to the less salient, rewarding outcome also present on that side. Consequently they focus on the punishment at the expense of the reward, and bias their responding away from the punished side. Thus, a failure to shift attention by either inadequate vIPFC recruitment or vIPFC inactivation disrupts the cost-benefit analysis by allowing the subject's choice to be unduly influenced by the negative outcome (3).

A similar explanation may account for the activation of human vIPFC that accompanies cognitive re-appraisal of emotional stimuli (23) (see also 24). Cognitive re-appraisal of a negatively-valenced picture requires subjects to shift their attention to potentially less salient,

but more positive interpretations of the picture, in order to diminish its negative affect. Thus, both cost-benefit decision making, as studied here, and cognitive re-appraisal may depend upon the attentional control functions of the vIPFC for the re-interpretation of emotional stimuli (25). Note that such re-interpretation is distinct from the situation that occurs in extinction, in which a negatively-valenced stimulus loses its negative valence, the latter having been shown to recruit primarily, ventromedial regions of PFC (26, 27). In the current task, the aversive stimulus does not lose its aversiveness, but instead we suggest that vIPFC-mediated attentional shifting facilitates the re-interpretation or re-evaluation of the aversive stimulus in the context of the competing reward.

It is unknown how the attentional-shifting role of the vIPFC is integrated into the decision making circuitry. The vIPFC region targeted here (area 12 including 12l, 12m and 12o) sends projections to both the mPFC and OFC (28, 29) but the relative contribution of these projections is unknown. However, while mPFC manipulations have been shown to cause immediate alterations in cost-benefit decision making (6), inactivation of the OFC in the present study did not, making it unlikely that interactions between vIPFC and antOFC were contributing to performance. Instead, projections from the vIPFC to the mPFC may provide the appropriate attentional bias, optimising the mPFC-mediated decision making process.

In contrast to the vIPFC, antOFC inactivation had no effect on the cost-benefit decision between primary reward and punishment but did result in a marked bias away from the punished side the following day, in the absence of punishment. This effect was blocked by inactivation of either the anterior hippocampus, amygdala or disconnection of the two, revealing the important role of the antOFC in moderating the formation of a punishment

memory, the expression of which is dependent upon an amygdala-anterior hippocampal circuit. Whether the punishment memory is also stored within this circuit remains to be determined. The OFC is implicated in the learning of stimulus-outcome relationships (9, 30). By learning to predict the presence of punishment, particularly in an environment in which punishment is sporadic, as in the present task, the antOFC may act to lessen the uncertainty about, and consequently the impact of, punishment, especially since uncertainty is a critical factor in the generation of anxiety (31). Of particular importance in the present study is the finding that despite intact decision making in the presence of punishment, inactivation of the antOFC caused aberrant consolidation of the emotionally arousing stimuli, leading to abnormal avoidance behaviour when confronted with a similar situation the subsequent day. Although the increased avoidance behavior seen after OFC inactivation rapidly returned to normal when the OFC was no longer inactivated, an OFC that is functionally compromised for longer may have deleterious psychopathological consequences for the amygdala and hippocampus, such as the potentiated punishment sensitivity seen in anxiety conditions. Such a proposal would provide a functional counterpart for findings that ventromedial (including OFC) PFC thickness correlates negatively with both amygdala/hippocampal activity and measures of anxiety (32, 33).

The amygdala and hippocampus are associated independently with anxiety, and anxiolysis (34), and increased grey matter density or activity in both structures is associated with increased sensitivity to negative stimuli (35). Both structures also have well established roles in the enhancement of memory for emotionally arousing events (36), which is consistent with the current findings that inactivation of either structure independently abolished the expression of the negative bias induced by OFC inactivation. However, the effects of crossed disconnection of the two emphasises the importance of their interaction in the expression of

this negative bias, which is consistent with the increased amygdala-hippocampal connectivity associated with trait sensitivity to aversive events and neuroticism (37, 38). The strength of amygdala-hippocampus connectivity has also been shown to increase bidirectionally during the retrieval of emotional information that is relevant to current behaviour, confirming the importance of their interaction (39, 40). However, until now the role of amygdala-hippocampal communication has been considered predominantly in the context of amygdala-mPFC connectivity and fear regulation (37, 41). The current findings reveal the importance of the OFC for modulating plasticity within this circuit in the regulation of anxious behaviour. Future studies will determine how the vIPFC interacts with subcortical structures such as the hippocampus and amygdala to regulate emotional decision making.

To conclude, we have shown that dysregulation of vIPFC and antOFC and their subcortical connections induce distinct patterns of punishment bias in an approach-avoidance conflict task. It is proposed that these biases are caused, respectively, by deficits in attentional control and punishment prediction. The differentiation of the component neural mechanisms underlying punishment processing revealed in the present study provides important new insight into the heterogeneity of mood and anxiety disorders, increasing our ability to predict the efficacy of specific treatment strategies in individual patients. For example, based on the present results, cognitive re-appraisal, a common component of Cognitive-Behavioural therapy, may be more successful in a patient poor at predicting, than one deficient in attentional control. These findings also highlight the importance of primate models for translational research, as such a detailed anatomical segregation of function between antOFC and vIPFC would not have been possible in species that do not share prefrontal homology with humans. Altogether, these findings advance our understanding of the prefrontal

organisation of executive functions, demonstrating the distinct role of prefrontal regions in the decision making process.

Materials and Methods

Subjects and Housing

Eight common marmosets (*Callithrix jacchus*; 3 females, 5 males) bred on site at the University of Cambridge Marmoset Breeding Colony were housed in pairs. All monkeys were fed 20 g of MP.E1 primate diet (Special Diet Services, Withams, Essex, UK) and two pieces of carrot five days a week after the daily behavioral testing session, with simultaneous free access to water for two hours. At weekends, their diet was supplemented with fruit, rusk, malt loaf, eggs, bread and treats and they had free access to water. Their cages contained a variety of environmental enrichment aids that were varied regularly and all procedures were performed in accordance with the UK Animals (Scientific Procedures) Act 1986.

Apparatus

Behavioral testing took place within a sound-attenuated box in a dark room. The animal sat in a clear, plastic transport box, one side of which was removed to allow the marmoset to reach through an array of vertical metal bars to touch stimuli presented on a touch-sensitive computer monitor (Campden Instruments, UK). When appropriate a reward of cooled banana milkshake (Nestlé, York, UK) was delivered to a centrally placed spout for 5 s, or a brief mildly aversive loud noise (0.3 s, 117dB) was played from a siren located centrally at the back of the test chamber. The test chamber was lit with a 3 W bulb. The stimuli presented on the monitor were green circles (40 mm diameter) with a small black dot in the middle (5mm diameter) which were displayed to the left and right of the central spout via the Whisker control system (42).

Behavioral Training and Testing

All monkeys were trained initially to enter a clear plastic transport box for marshmallow reward and accustomed to the test apparatus. Monkeys were then familiarized with the milkshake reward, and taught to respond to the touchscreen for reward until they were reliably and accurately making 30 responses or more to a green square presented to the left or right of the lickert in 20 minutes (for full experimental details, see 43). The stimuli were then changed to green circles and a variable interval (VI) schedule introduced gradually until the monkeys were happily responding to both stimuli equally. Reward was presented to each stimulus on an independent VI schedule, each with a mean schedule of 15 s (ranging from 5 to 25s in increments of 5s). All monkeys made approximately the same number for responses to both sides of the screen. If a response was rewarded, the stimulus remained on the screen for the duration of reward (5s). The aversive noise was then introduced. Initial presentations were at 90dB, which incremented gradually up to 117 dB with little or no deleterious effects on performance. If a response was unrewarded, the stimuli disappeared and then immediately reappeared to signal the start of the next trial. If a response was punished the stimuli disappeared, the aversive noise sounded (0.3 s) and the stimuli immediately reappeared as before. The testing period was limited only by time, and the session length was 12 minutes. The monkeys made an average of 150 responses during this time (75 to the left and 75 to the right; see Figure 1b). Once trained, the monkeys received cannulation surgery.

Cannulation surgery

Subjects were premedicated with ketamine hydrochloride (Pharmacia and Upjohn, 0.05 ml of a 100 mg/ml solution, i.m.) and given a long-lasting prophylactic analgesic (Carprieve; 0.03 ml of 50 mg/ml carprofen, s.c.; Pfizer, Kent, UK). They were intubated and maintained on isoflurane gas anaesthetic (flow rate: 2.0–2.5% isoflurane in 0.3 l/min O₂; Novartis Animal

Health UK, Herts, UK) and placed in a stereotaxic frame modified for the marmoset (David Kopf, Tujanga, CA, USA). Anaesthesia was monitored clinically and by pulse oximetry with capnography. Cannulae (Plastics One, Roanoke, VA, USA) were implanted into the ventrolateral prefrontal cortex (double 6mm long cannulae, 1mm apart, anteroposterior [AP] + 17.25/18.25, lateromedial [LM] adjusted in situ to give depth greater than 3mm at 8°) or the anterior OFC (single 6mm long cannula, AP + 17, LM \pm 3) having been adjusted where necessary *in situ* according to cortical depth (28). In the antOFC-cannulated animals, amygdala (single 14 mm long cannula, AP + 9.3, LM \pm 5.6, ventral [V] + 5) and anterior hippocampal (double 15 mm long cannula, 1mm apart, AP + 6, LM \pm 5.75/7.75, V + 5) cannulae were added separately in an additional surgery. However one antOFC-cannulated animal broke his leg and was euthanased prior to this. Postoperatively, and when fully recovered, all monkeys were returned to their homecage and then received the analgesic meloxicam (0.1 ml of a 1.5 mg/ml oral suspension; Boehringer Ingelheim, Germany) for 3 days as well as 10 days of ‘weekend diet’ and water *ad libitum* to allow complete recovery before returning to testing. Cannulae were cleaned every week (and caps and cannula blockers changed) to ensure the cannula site remained free from infection.

Drug treatments

For all infusions the monkey was held gently by a researcher. For central infusions, the caps and cannula blockers were removed and the site cleaned with alcohol. The sterile injector was inserted into the cannula and saline or muscimol/baclofen (0.5 μ l of 0.1 mM muscimol/1.0 mM baclofen) infused at a rate of 0.25 μ l/min for 2 mins. Injectors were left in place for 1 minute to allow diffusion of liquid before being removed, clean caps and dummies applied, and the monkey returned to the homecage for 20 mins. See table S1 for infusion order. In all cases the punishment sessions were intermittently interspersed between reward sessions, and occurred approximately once per week. The punishment was always presented on the side on

which the monkey responded to most the preceding day. Thus the location of punishment varied both between monkey, and between individual testing sessions for a given monkey. Infusions usually occurred twice a week (one saline treatment and one drug treatment) in a randomised order between groups. There was one session per subject per condition. For peripheral injections the site of injection was cleaned with alcohol, and then injected with either diazepam (0.25mg/kg, i.m, Wockhardt Ltd, UK) or an equal volume of saline 30 mins prior to testing.

Analysis

For each session, a bias measure was calculated for each animal. This was the ratio of the number of responses made to the non-preferred side over the number of responses made to the preferred side (bias = non-preferred responses/ preferred responses). To calculate the change in bias that resulted from any drug treatment, the drug induced bias was calculated as a percentage of the equivalent saline treatment. Thus a bias measure of 100% indicates that the effect of drug treatment was identical to that of saline treatment. A bias of above 100% indicates that the drug treatment induced a bias towards the non-preferred side i.e. away from the preferred (often punished) side, and a bias of less than 100% indicated a bias away from the non-preferred side and towards the preferred (and often punished) side.

Post mortem lesion assessment

All monkeys were premedicated with ketamine hydrochloride (Pharmacia and Upjohn, 0.05 ml of a 100 mg/ml solution, i.m.) and humanely euthanased with Euthatal (1ml of a 200 mg/ml solution, pentobarbital sodium; Merial Animal Health Ltd; i.p.) before being perfused transcardially with 500ml of 0.1M PBS, followed by 500ml of 4% paraformaldehyde fixative over approximately 15 mins. The entire brain was then removed and placed in further paraformaldehyde overnight before being transferred to a 30% sucrose solution for at least 48

hours. For verification of cannulae placement, coronal sections (60 μm) of the brain were cut using a freezing microtome, the cell bodies were stained using Cresyl Fast Violet and the sections viewed under a Leitz DMRD microscope. For each animal, cannula locations were schematized onto drawings of standard marmoset brain coronal sections and composite diagrams were then made to illustrate the extent of overlap between animals.

Statistics

Behavioral data were analysed using SPSS v.21 (IBM, NY, USA). For analysis of variance (ANOVA), homogeneity of variance was verified using Levene's test; type III sums of squares and full factorial models were used unless stated. Where applicable, the Huynh–Feldt correction was used to correct for any violations of the sphericity assumption as assessed by the Greenhouse–Geisser test. For statistical purposes, this data was square-root transformed to normalise its distribution according to Levene's test. However for clarity, the data presented in the figures is not transformed.

Acknowledgements

This research was funded by an MRC Programme Grant to ACR and a Career Development Award to HFC. It was conducted at the Behavioural and Clinical Neuroscience Institute, which is supported by a joint award from the Medical Research Council and Wellcome Trust (G00001354). All authors contributed extensively to the work presented in this paper and we thank Rudolf Cardinal for helpful advice and discussion and Mercedes Arroyo for histology. There are no competing financial interests.

References

1. Murphy FC, Michael A, Robbins TW, Sahakian BJ (2003) Neuropsychological impairment in patients with major depressive disorder: the effects of feedback on task performance. *Psychol Med* 33(3):455–467.
2. Dickson JM (2006) Perceived consequences underlying approach goals and avoidance goals in relation to anxiety. *Pers Individ Dif* 41:1527–1538.

3. Gold AL, Morey RA, McCarthy G (2014) Amygdala-Prefrontal Cortex Functional Connectivity During Threat-Induced Anxiety and Goal Distraction. *Biol Psychiatry*.
4. Mochcovitch MD, da Rocha Freire RC, Garcia RF, Nardi AE (2014) A systematic review of fMRI studies in generalized anxiety disorder: Evaluating its neural and cognitive basis. *J Affect Disord* 167C:336–342.
5. Adhikari A (2014) Distributed circuits underlying anxiety. *Front Behav Neurosci* 8:112. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24744710>.
6. Amemori K, Graybiel AM (2012) Localized microstimulation of primate pregenual cingulate cortex induces negative decision-making. *Nat Neurosci* 15(5):776–785.
7. Wittchen HU, et al. (2011) The size and burden of mental disorders and other disorders of the brain in Europe 2010. *Eur Neuropsychopharmacol* 21(9):655–679.
8. Noonan MP, et al. (2010) Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proc Natl Acad Sci U S A* 107(47):20547–20552.
9. Izquierdo A, Suda RK, Murray EA (2004) Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *J Neurosci* 24(34):7540–7548.
10. Kazama AM, Davis M, Bachevalier J (2014) Neonatal lesions of orbital frontal areas 11/13 in monkeys alter goal-directed behavior but spare fear conditioning and safety signal learning. *Front Neurosci* 8:37.
11. Izquierdo A, Suda RK, Murray EA (2005) Comparison of the effects of bilateral orbital prefrontal cortex lesions and amygdala lesions on emotional responses in rhesus monkeys. *J Neurosci* 25(37):8534–8542.
12. Fox AS, et al. (2010) Orbitofrontal cortex lesions alter anxiety-related activity in the primate bed nucleus of stria terminalis. *J Neurosci* 30(20):7023–7027. A
13. Kalin NH, Shelton SE, Davidson RJ (2007) Role of the primate orbitofrontal cortex in mediating anxious temperament. *Biol Psychiatry* 62(10):1134–1139.
14. Agustin-Pavon C, et al. (2012) Lesions of ventrolateral prefrontal or anterior orbitofrontal cortex in primates heighten negative emotion. *Biol Psychiatry* 72(4):266–272.
15. West EA, DesJardin JT, Gale K, Malkova L (2011) Transient inactivation of orbitofrontal cortex blocks reinforcer devaluation in macaques. *J Neurosci* 31(42):15128–15135.
16. Fuchs RA, Evans KA, Parker MP, See RE (2004) Differential involvement of orbitofrontal cortex subregions in conditioned cue-induced and cocaine-primed reinstatement of cocaine seeking in rats. *J Neurosci* 24(29):6600–6610.
17. Barbas H (2007) Flow of information for emotions through temporal and orbitofrontal pathways. *J Anat* 211(2):237–249.
18. Baxter MG, Gaffan D, Kyriazis DA, Mitchell AS (2009) Ventrolateral prefrontal cortex is required for performance of a strategy implementation task but not reinforcer devaluation effects in rhesus monkeys. *Eur J Neurosci* 29(10):2049–2059.
19. Bussey TJ, Wise SP, Murray EA (2001) The role of ventral and orbital prefrontal cortex in conditional visuomotor learning and strategy use in rhesus monkeys (*Macaca mulatta*). *Behav Neurosci* 115(5):971–982.

20. Dias R, Robbins TW, Roberts AC (1996) Primate analogue of the Wisconsin Card Sorting Test: effects of excitotoxic lesions of the prefrontal cortex in the marmoset. *Behav Neurosci* 110(5):872–886.
21. Rich ED, Wallis JD (2014) Medial-lateral organisation of the orbitofrontal cortex. *J Cogn Neurosci* 26(7):1347–1362.
22. Kennerley SW, Wallis JD (2009) Reward-dependent modulation of working memory in lateral prefrontal cortex. *J Neurosci* 29(10):3259–3270.
23. Buhle JT, et al. (2013) Cognitive Reappraisal of Emotion: A Meta-Analysis of Human Neuroimaging Studies. *Cereb Cortex*.
24. Riva P, Romero Lauro LJ, Dewall CN, Bushman BJ (2012) Buffer the pain away: stimulating the right ventrolateral prefrontal cortex reduces pain following social exclusion. *Psychol Sci* 23(12):1473–1475.
25. Wager TD, Davidson ML, Hughes BL, Lindquist MA, Ochsner KN (2008) Prefrontal-subcortical pathways mediating successful emotion regulation. *Neuron* 59(6):1037–1050.
26. Phelps EA, Delgado MR, Nearing KI, LeDoux JE (2004) Extinction learning in humans: role of the amygdala and vmPFC. *Neuron* 43(6):897–905.
27. Milad MR, et al. (2007) Recall of fear extinction in humans activates the ventromedial prefrontal cortex and hippocampus in concert. *Biol Psychiatry* 62(5):446–454.
28. Roberts AC, et al. (2007) Forebrain connectivity of the prefrontal cortex in the marmoset monkey (*Callithrix jacchus*): an anterograde and retrograde tract-tracing study. *J Comp Neurol* 502(1):86–112.
29. Vogt BA, Pandya DN, Rosene DL (1987) Cingulate cortex of the rhesus monkey: I. Cytoarchitecture and thalamic afferents. *J Comp Neurol* 262(2):256–270.
30. Schoenbaum G, Saddoris MP, Stalnaker TA (2007) Reconciling the roles of orbitofrontal cortex in reversal learning and the encoding of outcome expectancies. *Ann N Y Acad Sci* 1121:320–335.
31. Gray JA, McNaughton N (2000) *The Neuropsychology of Anxiety* (Oxford University Press, Oxford).
32. Foland-Ross LC, et al. (2010) Amygdala reactivity in healthy adults is correlated with prefrontal cortical thickness. *J Neurosci* 30(49):16673–16678.
33. Bishop SJ, Duncan J, Lawrence AD (2004) State anxiety modulation of the amygdala response to unattended threat-related stimuli. *J Neurosci* 24(46):10364–10368.
34. Barkus C, et al. (2010) Hippocampal NMDA receptors and anxiety: at the interface between cognition and emotion. *Eur J Pharmacol* 626(1):49–56.
35. Barros-Loscertales A, et al. (2006) Striatum gray matter reduction in males with an overactive behavioral activation system. *Eur J Neurosci* 24(7):2071–2074.
36. Roozendaal B, McGaugh JL (2011) Memory modulation. *Behav Neurosci* 125(6):797–824.
37. Tzschoppe J, et al. (2014) Aversive learning in adolescents: modulation by amygdala-prefrontal and amygdala-hippocampal connectivity and neuroticism. *Neuropsychopharmacology* 39(4):875–884.
38. Hahn T, et al. (2010) Functional amygdala-hippocampus connectivity during anticipation of aversive events is associated with Gray's trait "sensitivity to punishment." *Biol Psychiatry* 68(5):459–464.

39. Smith AP, Stephan KE, Rugg MD, Dolan RJ (2006) Task and content modulate amygdala-hippocampal connectivity in emotional retrieval. *Neuron* 49(4):631–638.
40. Fastenrath M, et al. (2014) Dynamic Modulation of Amygdala–Hippocampal Connectivity by Emotional Arousal. *J Neurosci* 34(42):13935–13947.
41. Kim MJ, et al. (2011) The structural and functional connectivity of the amygdala: from normal emotion to pathological anxiety. *Behav Brain Res* 223(2):403–410.
42. Cardinal RN, Aitken MR (2010) Whisker: a client-server high-performance multimedia research control system. *Behav Res Methods* 42(4):1059–1071.
43. Roberts AC, Robbins TW, Everitt BJ (1988) The effects of intradimensional and extradimensional shifts on visual discrimination learning in humans and non-human primates. *Q J Exp Psychol [B]* 40(4):321–341.
44. Burman KJ, Rosa MG (2009) Architectural subdivisions of medial and orbital frontal cortices in the marmoset monkey (*Callithrix jacchus*). *J Comp Neurol* 514(1):11–29.

Figure Legends

Figure 1. Behavioral task and prefrontal cannulae placements.

A. Responding to either of two identical visual stimuli presented to the left and right of a touch sensitive computer screen gained reward (5s banana juice) according to independent but identical variable-interval schedules. In occasional test sessions (average one per week), responding to one of the stimuli also resulted in punishment (0.3s 117dB loud noise) on a leaner independent variable-interval schedule, while the reward schedule was unchanged.

B. Under control conditions marmosets responded relatively equally to both stimuli, with only a slight preference for one side. Accordingly that side received the punishment on punishment sessions and thus, overall, there is more responding to the ‘punished’ or ‘to be’ punished (P) side than the ‘non-punished’ or ‘to be’ non-punished (NP) side ($F_{1,7} = 26.08$, $P = 0.001$). This remained the same regardless of whether punishment was present or not, ($F_{1,7} =$

0.16, $P = 0.698$; see text for detailed explanation of why animals in control conditions did not avoid punishment).

C. Sagittal marmoset MRI section illustrating the rostro-caudal locations of the vIPFC and OFC for target infusions.

D. Schematics showing the single and double intracerebral cannulae targeting respectively, area 11/antOFC and area 12/vIPFC, together with the actual cannulae locations for each animal and representative histological sections with arrows marking the position of the cannulae. All cannulae were located within the range of AP 15.8-16.6, plotted here on a single coronal section for each target area. Scale bar = 5mm. Cytoarchitectonic parcellation is according to Burman and Rosa, 2009 (44) and the circles represent the estimated maximal spread of the muscimol/baclofen or saline infusions (15).

Figure 2. Inactivation of either the vIPFC or antOFC induces a negative decision bias, but on different timescales

A. Inactivation of either the vIPFC or OFC did not affect responding when only reward was present (left two pairs of bars), but produced differential effects on responding when punishment was introduced (right two pairs of bars). vIPFC inactivation (green bars) caused a bias away from punishment on the day of punishment ('infusion' day), while OFC inactivation (blue bars) caused a bias away from punishment the day after ('next' day). B. The overall number of responses was not affected by either inactivation. The region of inactivation ('vIPFC' or 'OFC'), the day of inactivation ('infusion' day or the 'next' day) and the presence of reward (blue droplet) and/or punishment (bell) are all indicated in the grid below the bars. A response bias of 100% indicates that an inactivation was identical to that of

saline treatment. See Materials and Methods for details of bias calculation.* $P < 0.05$ on square root transformed data. Data are represented as mean \pm SEM

Figure 3. The negative decision biases were abolished by anxiolytic treatment

A. The bias away from the punished side that is seen after vIPFC inactivation on a punished day (left dark green bar) was abolished by the concomitant presence of Diazepam (left light green bar). Next day performance was unaffected (right green bars). B. The bias away from the punished side that is seen on the next day after OFC inactivation on a punished day (right dark blue bar) was completely abolished when Diazepam was administered on the infusion day (right mid-blue bar) and partially abolished when Diazepam was administered on the next day (right pale blue bar). $P < 0.05$ on square root transformed data. The region of inactivation ('vIPFC' or 'OFC'), the day of inactivation ('infusion' day or the 'next' day) and the presence of diazepam ('D'), reward (blue droplet) and punishment (bell) are all indicated in the grid below the bars. Data are represented as mean \pm SEM.

Figure 4. AntOFC inactivation modulates a punishment memory within an amygdala-hippocampal circuit

A. Sagittal marmoset MRI section illustrating the rostro-caudal coordinates of the antOFC, amygdala and anterior hippocampus for cannula placements (n=3).

B. The effects of amygdala, and anterior hippocampal inactivation or their crossed disconnection on the next day, after antOFC inactivation on a punished day, were compared to the effects of amygdala/hippocampal manipulations on the next day, after punishment in the absence of antOFC inactivation. The bias away from the punished side that is seen on the

next day after OFC inactivation on a punished day (right dark blue bar compared to left dark blue bar) was completely abolished if the anterior hippocampus (right red bar), amygdala (right khaki bar) or a unilateral crossed disconnection of both (right purple bar) were **also** inactivated the next day. * $P < 0.05$ on square root transformed data . The region of inactivation (antOFC; 'OFC', amygdala; 'Amyg', hippocampus; 'Hipp', or hippocampal/amygdala disconnection; 'H/A'), the day of inactivation ('infusion' day or the 'next' day) and the presence of reward (blue droplet) and punishment (bell) are all indicated in the grid below the bars. Data are represented as mean \pm SEM.

C and D. Schematics illustrating the location of the amygdala and anterior hippocampal cannulae in a coronal section for each animal, alongside representative histological sections with arrows marking the position of the cannulae tracts.