Inactivation of the ventrolateral orbitofrontal cortex impairs flexible use of safety signals

Mary C. Sarlitto, **Allison R. Foilb**, and **John P. Christianson**[*]

Department of Psychology, Boston College, Chestnut Hill, MA 02467, USA

Abstract

Survival depends on adaptation to shifting environmental risks and opportunities. Regarding risks, the mechanisms which permit acquisition, recall, and flexible use of aversive associations is poorly understood. Drawing on the evidence that the orbital frontal cortex is critical to integrating outcome expectancies with flexible appetitive behavioral responses, we hypothesized that OFC would contribute to behavioral flexibility within an aversive learning domain. We introduce a fear conditioning procedure in which adult male rats were presented with shock-paired conditioned stimulus (CS+) or a safety cue (CS−). In a recall test, rats exhibit greater freezing to the CS+ than the CS−. Temporary inactivation of the ventrolateral OFC with muscimol prior to conditioning did not affect later discrimination, but inactivation after learning and prior to recall impaired discrimination between safety and danger cues. This result complements prior research in the appetitive domain and suggests that the OFC plays a general role in behavioral flexibility regardless of the valence of the CS.

INTRODUCTION

An animal's prosperity and survival require flexible adaptation to a constantly changing environment. Past experience shapes decision making in part through the accumulation of learned associations between stimuli and their associated outcomes, which can be used to make predictions and decisions about future behaviors. Areas in the frontal cortex have been especially implicated in this type of associative learning and, of these, the orbital frontal cortex (OFC) seems to be important in mediating cognitive flexibility (Dalley et al., 2004; Murray et al., 2007; Rudebeck and Murray, 2014). There is a vast and rapidly growing body of literature in which the OFC appears to contribute to a wide array of cognitive functions,

[*]Corresponding Author: John P. Christianson, Gianinno Family Sesquicentennial Assistant Professor, Department of Psychology, Boston College, McGuinn 300, Chestnut Hill, MA 02467, P: 1-617-552-3970, j.christianson@bc.edu.

from novel associative learning, reward valuation, reversal learning, and extinction (Stalnaker et al., 2015).

With the goal of resolving complexities regarding the role of OFC in cognitive flexibility, the 'cognitive map' hypothesis accounts for much of the empirical data regarding OFC function (Wilson et al., 2014). A cognitive map is the mental representation of the external environment (Gallistel, 1989). This theory proposes that the OFC maintains a cognitive map characterized by the current task state. For example, using a Pavlovian conditioning paradigm, a rat might be trained with repeated cue presentations to associate a conditioned stimulus (CS, like a flashing light) with the occurrence of an unconditioned stimulus reward (US, food pellet). As the rat begins to behaviorally distinguish between CS/no-CS trials, a reflection of differential reward expectancy, the cognitive map theory suggests that the two conditions become encoded as separate 'task states'. Tracking task states, particularly those occurring in the same context but with opposing outcomes, is considered by this theory to be a key function of OFC activity (Wilson et al., 2014). Thus, loss of OFC function does not always impair state-relevant processing but does invoke noticeable impairments to performance of tasks requiring differentiation between perceptually similar states.

The cognitive map model captures a significant amount of data relating to the function of OFC but a limitation of the theory stems from the nearly exclusive reliance on behavioral tasks in which the outcome expectancies related to the value of desirable stimuli, such as food. Regarding aversively motivated behaviors, such as conditioned fear in which several contributions of the OFC are reported (*e.g.,* Rodriguez-Romaguera et al., 2015; Zimmerman et al., 2018) but there are conflicting results (Shiba, Santangelo, & Roberts, 2016). A test of OFC function in a learning context in which the conditioned stimuli become associated with different expectations of aversive outcomes would provide a systematic test of the current theory's generality and provide insight to the neural circuitry underlying cognitive flexibility in the face of environmental stimuli that predict danger. We recently reported that discriminative fear conditioning leads to differential freezing upon presentations of conditioned danger (shock paired CS+) and conditioned safety (unpaired CS−) cues (Chen et al., 2016; Foilb et al., 2016). In the danger and safety learning scenario, the CS+ and CS− could be argued to reflect distinguishable task states that can be called upon to guide behavioral responses, and therefore involve the OFC.

The OFC receives polymodal sensory input and projects extensively to the limbic system and midbrain motor regions capable of governing behavioral output (Ongür and Price, 2000; Price, 2007). It is this intersection of polymodal sensory input and visceromotor outputs that makes the OFC relevant to reward processing but also positions it well to map the outcome expectancies of undesirable aversive stimuli. The rat OFC is subdivided into the medial (MO), ventral (VO), ventrolateral (VLO) and lateral (LO) and anterior agranular insular cortex (AI; Price, 2007; Price, 2007). Anatomical, mechanistic and recording experiments targeting these regions provide evidence for some regional functional specificity with regard to decision-making, with medial regions relating more to affective regulation with lateral regions more to sensory integration (Rempel-Clower, 2007; Izquierdo, 2017). The vast majority of prior research supporting the cognitive map theory entailed manipulations or electrode recordings from the LO and AI region (Murray et al., 2007; Rudebeck and Murray,

2014; Wilson et al., 2014). Here we use a fear discrimination test with multiple cue trials presented in a quasi-random order such that freezing, the behavioral expression of fear, appears to be under the flexible control of the danger or safe signals. This behavioral flexibility should entail the formation and use of a cognitive map and so provides a paradigm suitable to assess the generality of this role of the OFC in both learning and flexible control of behavior by aversive cues. The AI does not seem to be required for this task (Foilb et al., 2016), therefore we targeted the LO/VO regions with pharmacological inactivation in fear discrimination acquisition or later recall.

EXPERIMENTAL PROCEDURES

Subjects

Forty-eight adult male Sprague-Dawley rats from Taconic (Hudson, NY), weighing 250–300g upon arrival were used. Before surgery, all rats were housed 2 per cage and housed singly after surgery with a short piece of autoclaved manzanita wood for enrichment. All animals were maintained on a 12-hour light/dark cycle within the Boston College Animal Care Facility and allowed to habituate to their home cages for one week before the start of surgery or behavior. All experimental protocols were reviewed and approved by the Boston College Institutional Animal Care and Use Committee (IACUC). The fear discrimination acquisition and recall data from 8 rats in Experiment 1 receiving polymodal cues were included in a large N analysis of this fear conditioning paradigm published previously (Foilb et al., 2018).

Surgical Implantation of Microinjeciton Cannula

As previously (Foilb et al., 2016), rats were anesthetized using isofluorane (3% in O_2) and mounted in a stereotaxic apparatus. An incision was made in the center of the scalp to expose bregma and lambda. Stainless steel guide cannula (22g; Plastics One, Roanoke, VA) were implanted bilaterally within the OFC (coordinates: +3.2mm anterior/posterior from bregma, ±2.2mm medial/lateral from midline, −3.4mm dorsal/ventral from dura) according to the atlas of Paxinos and Watson (George and Charles, 2013) to target the VO/LO area. Cannulas were fixed to the skull using stainless steel screws and acrylic cement. A stylet extending 1mm ventral to the tip of the cannula was guided into each side and tightened to the top of the fixture to ensure patency. After surgery, each rat received 1 dose each of loxicom (1mg/kg), penicillin G procaine (15,000 Units), and 5mL of lactated Ringers' solution (Henry Schein, Albany, NY) to aide in recovery. The next day, rats were administered a second dose of loxicom in accordance with the policy of the Boston College IACUC. All animals were allowed one week of post-operative recovery before the start of behavioral testing. During the recovery period, each rat was periodically handled and stylets checked to acclimate the animals to this type of contact, and also to confirm that cannulas remained clear.

Fear Discrimination Conditioning

Conditioning occurred as previously (Chen et al., 2016; Foilb et al., 2016; Foilb and Christianson, 2016; Foilb et al., 2018) in black plastic boxes, $10 \times 11 \times 6$-in (L × W × H) with shock grids (Model H10-11R-TC-SF, Coulbourn Intruments, Whitehall, PA), and wire

mesh tops placed inside a ventilated light and sound-attenuating enclosure, $15 \times 12 \times 27$ in ($L \times W \times H$). Two commercially available arrays of infrared LED lights (CMVision Model IR30) illuminated the chamber. Behavior was recorded using overhead cameras (Model VX-5000, Microsoft, Redmond, VA) with infrared blocking filters replaced by infrared passing filters. Freezing was quantified by computer (ANY-Maze version 4.99, Stoelting, Wood Dale, IL), using the settings recommended by the manufacturer (as previously verified in Christianson et al., 2011). A white LED array (Model LPL620WTHD, Hampton Bay) and a speaker mounted at the top of the chamber were used for delivery of conditioned stimuli.

Conditioning entailed 15 presentations of each the CS+ (the danger cue) and the CS− (the safety signal). A white noise pip sound (pip-duration = 10 ms, interval = 3 Hz, 75 dB), and flashing LED lights (264.0 Lux, on/off, 20Hz) were used as the conditioning cues. In one experiment these cues were compared to distinguishable auditory cues (described below). Assignment of light or pip as the CS+ or CS− was counterbalanced in each experiment. Each conditioning session began with 2 minutes of context exposure to habituate to the chamber alone. Conditioning trials began with a 5 s, 1kHz tone, followed immediately by a 15s presentation of either the CS+ or CS−. The trials were presented in a quasi-random order such that no trial type was presented more than twice in succession. CS+ trials co-terminated with a 500 ms, 1.2 mA shock and CS− trials did not. A 70 s inter-trial interval followed the completion of the cue presentation before the next trial began. Recall tests were performed in the same apparatus and consisted of a 2 min exposure to the context followed by 10 CS+ and 10 CS− (30 s duration) presented in quasi-random order with a 30s inter-cue-interval.

Pharmacological inactivation of OFC

Microinjections of the GABA$_A$ agonist muscimol were made 1 h prior to either conditioning or recall tests as previously (Foilb et al., 2016). Muscimol was selected for its capacity to hyperpolarize cells through agonism of the GABA$_A$ chloride channel, and thus inhibit local neuronal activity in the region of interest (Majchrzak and Di Scala, 2000). Bilateral infusions of 50 ng muscimol (500 nL of 100 ng/μL musimol in 0.9% saline, Sigma, St. Louis, MO) or saline alone at a rate of 1 μL/min. Injectors (33g, Plastics One) were left in place for an additional minute to allow for diffusion. At the end of each experiment rats were overdosed with tribromoethanol. Brains were dissected and immediately flash frozen in 2-methylbutane on dry ice and stored in a −80°C freezer. 40μm thick coronal slices were collected of the OFC, mounted onto gelatin-subbed slides, stained with cresyl violet, washed with histoclear, and allowed to dry before examination under a brightfield microscope. Cannula placements in the vlOFC were determined through comparison with the *Rat Brain Atlas in Stereotaxic Coordinates* (Paxinos & Watson, 2013).

Experiment 1: Effect of CS modality on fear discrimination—A key aim in our prior work was to investigate the interaction of CS+ and CS− when presented in compound, a goal that favored the use of CSs of different sensory modality (i.e. a light and a sound). However, much of the prior work regarding the contribution of the vlOFC has used only auditory cues. In establishing the behavioral protocol to test the function of vlOFC in fear discrimination learning and recall, we first compared our conditioning procedure, as in (Foilb et al., 2016), using multimodal cues to the same protocol using polyphonic auditory

CSs as in (Wright et al., 2015). On Day 1 all rats (N=16) were trained using the CS+/CS− conditioning program – half with the multimodal cues (light vs. pip) and half with the unimodal cues (a stapler sound vs. car horn). On Day 2 freezing behavior was quantified during recall tests. As previously (Foilb and Christianson, 2016), we repeated this cycle of conditioning and recall for 3 additional days to establish a learning curve. Cue modality did not appear to influence behavior as freezing to the CSs during conditioning and recall were equivocal when comparing between multimodal and unimodal conditions (Figure 1). We selected the multimodal CS preparation for the later inactivation experiments to be consistent with our prior work.

Experiment 2: Effect of pre-conditioning OFC inactivation on fear discrimination—To determine if the OFC contributes to the initial learning of the fear discrimination, 16 rats implanted with bilateral OFC cannula were given either muscimol or saline injections 1h prior to fear discrimination conditioning with the light and pip CSs. Behavioral freezing was quantified during the conditioning session and in a drug-free recall test 24 h later.

Experiment 3: Effect of post-conditioning OFC inactivation on fear discrimination during recall—To determine if the OFC is involved in the behavioral responses to danger (CS+) and safety (CS−) 16 rats implanted with bilateral OFC cannula were given 3 consecutive days of fear discrimination conditioning, with a recall test each following morning. Three conditioning sessions were used to ensure that all rats exhibited differential freezing to the CS+ and CS− during recall tests. 24h after the third conditioning session, one half of the rats received bilateral muscimol infusions 1h before a recall test and the other half received a saline injection. Behavioral freezing was quantified during the recall test and the rats were returned to their vivarium. To allow a within-subjects comparison, 24h later all rats received a second round of microinjections but drug treatment was switched such that rats that received muscimol in the first test, received saline in the second test and vice versa.

Data Analysis

Time spent freezing during CS presentation was converted to a percent time and averaged into 3 trial bins for conditioning and into 2 trial bins for recall tests. Data were found to fit normal distributions (Shapiro-Wilk Normality Test) and homogeneity of variance was tested with Bartlett's test. Sphericity was assumed for repeated measures designs. Analysis of variance (ANOVA) were used to examine the influence of the experimental variables on freezing. Post hoc tests were performed using the Sidak correction to control experiment-wise type one error to $\alpha = 0.05$ (Hsu, 1996). When only two experimental conditions were compared, paired or unpaired t-tests were used as noted.

RESULTS

Experiment 1

To assess fear discrimination learning and the possible influence of cue type we observed behavioral freezing over a series of 4 conditioning (Figure 1A) and recall tests (Figure 1B).

Each conditioning and recall test was analyzed by a 3-way ANOVA with Cue type (unimodal vs. polymodal) as a between-subjects variable and CS type (CS+ vs CS−) and Trial block as within-subjects variables. In conditioning there were significant main effects of Trial on Day 1, $F(4, 56) = 7.32, p < 0.001,$ and Day 2, $F(4, 56) = 3.23, p < 0.02,$ and a main effect of CS type on Day 4, $F(1, 14) = 8.33, p = 0.01.$ In recall there were significant main effects of CS on test Day 1, $F(1, 14) = 41.77, p < 0.001,$ Day 2, $F(1, 14) = 54.82, p < 0.001,$ Day 3, $F(1, 14) = 103.20, p < 0.001,$ and Day 4, $F(1, 14) = 97.92, p < 0.001;$ main effects of Trial on Day 1, $F(4, 56) = 12.09, p < 0.001,$ Day 2, $F(4, 56) = 15.23, p < 0.001,$ Day 3, $F(4, 56) = 13.27, p < 0.001,$ and Day 4, $F(4, 56) = 21.57, p < 0.001;$ and Trial by CS interaction on Day 1, $F(4, 56) = 5.57, p < 0.001,$ and Day 2, $F(4, 56) = 8.46, p < 0.001.$ No other main effect or interaction reached significance. Fear discrimination was evident as differential freezing to the CS− compared to the CS+ in all recall tests, but not during conditioning sessions. Regarding the role of stimulus modality on freezing and discrimination, in no cases did significant main effects of Cue type reach significance. Therefore, post hoc comparisons were made between the CS+ and CS− by Trial pooled across stimulus modality. During conditioning, there were significant effects of Trial on days 1 and 2 which reflect great freezing in trial block 2 than blocks 1 or 3 (ps < 0.036), but no other trial comparisons reach significance. There was a significant main effect of CS on day 4 indicating discrimination between the CS+ and CS−.

With regard to the recall tests, there were significant effects of CS and Trial on all test days which reflect clear behavioral discrimination between the CS+ and CS−. On days 1 and 2 there were also significant CS by Trial interactions, which reflect increasing differential freezing to the CS+ and CS−. Post hoc tests were made to compare the CS+ CS− at each trial; because there were no effects of stimulus type, the unimodal and polymodal data were pooled. Freezing to the CS− was significantly less than freezing to the CS+ on every trial across 4 days, ps < 0.001 except for Test 1 trials 1 and 2, and Test 2 trial 1 in which ps = 0.02, 0.01, and 0.018 respectively.

Experiment 2

To determine the role of the OFC in the acquisition of a conditioned fear discrimination, muscimol or vehicle was injected 60min prior to the conditioning phase on day 1, n = 8/group and freezing was observed during fear discrimination conditioning and a recall test 24h later (Cannula placements are depicted in Figure 2). Freezing during conditioning (Figure 3A) and recall phases (Figure 3B) was analyzed with separate 3-way ANOVAs with drug condition as a between-subjects factor, and CS type and Trial as within-subjects factors. In neither phase did the main effect of drug reach significance, conditioning: $F(1, 14) = 1.00, p = 0.36,$ recall, $F(1, 14) = 0.11, p = 0.75,$ and there were no significant interactions with drug (ps > 0.05). In the conditioning phase there was a significant main effect of Trial, $F(4, 56) = 2.91, p = 0.03,$ but no other effects or interactions reached significance; no post hoc comparisons between trials reached significance. In the discrimination recall test there was a significant main effect of Cue, $F(1, 14) = 85.91, p < 0.001,$ and a significant Trial by Cue interaction, $F(1, 14) = 6.12, p < 0.03.$ Post hoc comparisons revealed significantly less freezing to the CS− compared to the CS+ at each trial, ps < 0.001. To summarize discrimination behavior, and facilitate the display of

individual subjects, a discrimination index was computed as a ratio of total freezing to the CS− divided by freezing to the CS+ multiplied by 100 (Figure 3C). There was no difference between drug conditions, $t_{unpaired}(14) = 0.39, p = 0.70$.

Experiment 3

To determine whether OFC contributes the flexible use of learned danger and safety signals, rats received fear discrimination conditioning on days 1 and 2, with discrimination recall tests on the morning of day 2 and day 3 (Tests 1 and 2; Figure 4A). Two rats were excluded from analysis because of misplaced cannula resulting in $N = 14$. A discrimination index was computed on day 3 and rats were rank-ordered and then assigned in alternating order to either vehicle or muscimol treatments. On day 4 rats received drug injection 60 min before a fear discrimination test (Test 3) based on group assignment. On day 5, rats received a 4th discrimination test (Test 4) after receiving the opposite drug treatment for a within subjects control. In other words, rats that received muscimol in Test 3, were tested 24h later in Test 4 with vehicle (Figure 4B). As expected, there was robust differential freezing to the CS+ and CS− in the first and second discrimination recall tests (drug-free), which was confirmed by 2-way ANOVAS with CS type and Trial as within-subjects factors. In Tests 1 and 2 there were main effects of Trial, $F(4, 52) = 6.11, p < 0.001$ and $F(4, 52)=8.97, p < 0.001$, respectively, and main effects of Cue, $F(1, 13) = 52.26, p < 0.001$ and $F(1, 13) = 58.74, p < 0.001$, respectively. Post hoc comparisons revealed significantly different freezing between the CS+ and CS− on all trial blocks with ps < 0.004.

Muscimol prior to recall testing appeared to interfere with fear discrimination especially in the later trials of each test (Figure 4B). Data from Test 3 and Test 4 were analyzed with separate 3-way ANOVAs with Drug group as a between-subjects factor with CS type and Trial as within-subjects factors. In Test 3 there were significant main effects of Drug, $F(1,12)=5.28, p = 0.04$, CS type, $F(1,12)=103.02, p < 0.001$, Trial, $F(4, 48) = 2.73, p = 0.04$ and a significant CS type by Drug interaction, $F(1, 12) = 5.97, p = 0.03$. In Test 4 there were significant main effects of Drug, $F(1, 12) = 10.45, p = 0.007$, CS type, $F(1, 12) = 45.18, p < 0.001$ and a significant CS type by Trial interaction, $F(4, 48) = 8.34, p < 0.001$. To explore these effects, we compared freezing to the CSs by drug treatment, by trial (*e.g.* Trial 1, CS+, Muscimol vs. Vehicle) and we compared to freezing to CS type by drug treatment for each trial (*e.g.* Trial 1, Muscimol, CS+ vs. CS−). In Tests 3 and 4, There was significantly less freezing to the CS− in the vehicle group compared to muscimol on Trials 4 and 5, ps < 0.019. Discrimination was evident in both drug groups with CS+ differing from CS− on all trials in the vehicle condition in both Tests 3 and 4 (ps < 0.028) and in the muscimol condition in Test 3 trials 1, 2, 3, and 4 (ps < 0.012) and Test 4 trials 1, 2, 3, and 5 (ps < 0.027).

The pattern of behavior in Tests 3 and 4 was consistent except for a slight reduction in overall freezing in Test 4 as seen in Figure 4B, therefore we pooled data from Tests 3 and 4, averaged freezing across trials, and freezing was analyzed with a 2-way ANOVA with Cue type and Drug as within-subjects factors (Figure 4C). Here there was a significant main effect of CS type, $F(1,13) = 97.92, p < 0.001$, Drug, $F(1, 13) = 13.68, p = 0.002$, and CS type by Drug interaction, $F(1, 13) = 5.00, p = 0.043$. All pairwise comparisons were made and

revealed significant differences between all possible groups ($ps < 0.001$) except for CS+ muscimol vs. CS− muscimol which approached significance, $p = 0.056$. The apparent discrimination impairment after muscimol could reflect a generalized effect of muscimol on freezing *per se.* To isolate the discrimination component, discrimination ratios were computed for the pooled data in Figure 4C. The discrimination ratio was significantly greater after muscimol than after vehicle indicating weaker inhibition of freezing to the CS−, paired *$t(13) = 4.03$, $p = 0.001$* (Figure 4D). In visual inspection of freezing during Tests 3 and 4 (Figure 4B) it appeared that in the muscimol treated group, freezing to the CS− was initially similar to the vehicle group, but drifted toward the CS+ level over repeated trials. To quantify this trend, we computed discrimination ratios for the first and last trial blocks. The discrimination index was greatest in the muscimol trial 5 conditioning (Mean = 71.55) which was greater than vehicle at both trial 1 (Mean = 32.95) and trial 5 (Mean = 31.37) but not from muscimol trial 1 (Mean = 40.65). However, a 2-way ANOVA with Trial and Drug as within-subjects factors only detected a significant effect of Drug, *$F(1, 13) = 6.53$, $p = 0.024$*; the interaction of Trial and Drug did not reach significance, *$F(1, 13) = 4.05$, $p = 0.065$.* Thus, the trend that OFC muscimol effects occurred primarily in the test trials requires additional investigation.

DISCUSSION

To test the hypothesis that the OFC is involved in the flexible responses to cues signaling danger and safety we established a discriminative fear conditioning paradigm in which a shock paired CS+ led to robust behavioral freezing and an unpaired, safe CS− did not. In a recall test in which the CS+ and CS− cues were presented in a quasi-random, alternating series, rats modulated behavioral freezing to the CSs indicating discrimination. Using this paradigm, pharmacological inactivation of the VO/LO region of the OFC prior to conditioning did not appear to influence any aspect of the fear discrimination conditioning, or the later flexible response to the CSs. In contrast, when VO/LO was inactivated prior to a recall test, discrimination was impaired. The impairment was more prominent in the later test trials in which freezing responses to the CS− increased which suggests that OFC is not involved in the initial recall of the discrimination. These results are generally consistent with the literature implicating the OFC and its outputs in tasks requiring cognitive flexibility. Importantly, the demonstration of a role for OFC in flexible responses in an aversively motivated learning paradigm suggests that the "cognitive map" theory of OFC function generalizes to many types of learning.

Although the OFC contributes to attention and salience assignment (Kahnt and Tobler, 2013; Ogawa et al., 2013) which would suggest a role for the OFC in the initial danger learning and discrimination, there are a number of reports using appetitively-motivated behavioral paradigms in which OFC manipulations left related learning processes intact. For instance, OFC lesions or inactivation do not consistently interfere with simple Pavlovian conditioning (Gallagher et al., 1999; Gremel and Costa, 2013; McDannald et al., 2005; Ostlund and Balleine, 2007) or discrimination learning in which a CS is paired with a desired stimulus such as sucrose (Izquierdo et al., 2004; McDannald et al., 2005; Schoenbaum et al., 2002; Walton et al., 2010). With regard to danger learning process, in a discriminative contextual fear conditioning task, pre-training OFC lesions impaired later discrimination between a

shock paired chamber and an unpaired chamber (Zelinski et al., 2010) but this effect may be attributed to generalization (Trow et al., 2017). That OFC inhibition did not influence any aspect of fear discrimination learning is in agreement with the wealth of prior studies. In light of these data, we must ask: *If not in the VO/LO OFC, where does fear discrimination learning occur?* Neither ventral hippocampus (Chen et al., 2016), posterior insular (Foilb et al., 2016), prelimbic or infralimbic cortex inactivations impair fear discrimination, although infralimbic inactivation prevented conditioned inhibition (Sangha et al., 2014). In the amygdala, fear versus safety conditioning lead to a substantial population of CS− responsive single units within the basolateral amygdala (Sangha et al., 2013), differential cue evoked responses (Rogan et al., 2005), and altered the structure of excitatory synapses (Ostroff et al., 2010) in the lateral amygdala. Thus, the CS-US associations underlying the flexible response to the CS+ and CS− observed here are likely encoded in a circuit that includes the amygdala and its inputs which are relayed to the OFC (Lichtenberg et al., 2017) which may interact with the IL to regulate freezing expression.

In Experiment 3, pharmacological inhibition of the OFC both increased fear expression *per se* (Main effect of Drug, Figure 4C) and interfered with behavioral flexibility in the discrimination recall test (Figure 4D). This task involves sequential presentations of danger and safety cues, in an order such that the same cue is never presented more than twice in succession. That behavioral freezing levels reliably modulate according to the CS type suggests one of two possibilities: that the OFC plays a general role in the inhibition of freezing or that it tracks the changing expectations of danger. Interference with either function would manifest as a deficit in fear discrimination. A direct role of VO/LO region of the OFC in control of freezing is certainly plausible given the reciprocal connections between the orbital cortex and the amygdala (Ongür and Price, 2000). Interestingly, OFC inhibition did not influence overall fear levels or fear discrimination in the first few safety trials. One would expect that if fear inhibition was the prevailing contribution of OFC in the current task, then the discrimination impairment would be evident in the early recall trials. On the other hand, over the course of the test, which would involve toggling between relatively high and low freezing levels several times, OFC inhibition appeared to bias the behavioral response to the danger signal. Our finding is consistent with two procedurally different, but conceptually related, reports in which OFC lesion (Clarke et al., 2015) or chemogenetic inhibition (Zimmermann et al., 2018) did not acutely alter aversively motivated behavior (punishment and conditioned fear, respectively), but did increase the influence of the aversive CS on behavior in later tests (i.e. fear extinction recall). The cognitive map hypothesis proposes that the major function of the OFC is to encode and mediate the shifts between 'states' on a cognitive map of task space which is especially relevant when the available predictive cues are perceptually similar but conceptually disparate (Wilson et al., 2014; Wilson et al., 2014). Because the CSs are presented in same apparatus and at the same trial intervals, behavioral flexibility throughout the test would require frequent alterations between of shock expectation with the output of the fear circuitry. Importantly, the test itself contains many cues (CS+, odors, and context) that have excitatory associations with the shock US and so disruption of the OFC likely leads to a bias in which the most salient stimuli, i.e. those that predict danger, gain control over behavior even in the presence of well-established safety signal; which is consistent with the inference

proposed by Clarke et al. (2015). Thus, we speculate that the impairment in fear discrimination observed here may be a consequence of interference with this task switching function rather than a role in fear inhibition *per se*.

Heightened fear is a hallmark symptom of PTSD (Yehuda et al., 2015) and failure to utilize learned safety cues has been suggested as a biomarker of PTSD (Grillon and Morgan, 1999; Jenewein et al., 2016; Jovanovic et al., 2009; Jovanovic et al., 2012). What is striking about this phenomenon is that while individuals with PTSD understand the meaning of the safety cue, fear enhanced startle and other biomarkers of anxiety appear insensitive to regulation by the safety cue (Jovanovic et al., 2009). Similarly, in Experiment 3, it was clear that inactivation of the OFC did not impede the initial recall of the danger/safety discrimination but did impair the use of this information over time. Our results suggest that a consequence of reduced OFC volume and hypoactivity, which are well documented in PTSD (Hakamata et al., 2007; Liberzon and Martis, 2006), is an impairment in switching out of a fearful state even when well-established safety cues are available.

The present study established a paradigm that allowed us to quantify recall of fear discrimination between safety and danger cues by measuring behavioral freezing. We used this model to investigate the function of the VO/LO region of OFC during aversive conditioning, an area that is sparse within current OFC literature. The current results reinforce a prevailing view of OFC function in the flexible use of learned associations to modulate behavioral responses. Importantly, our results reveal an effect of OFC inactivation specifically when rats have had to endure numerous switches between danger and safe state. In this case, the danger response prevailed and this may explain why impairments in OFC function are commonly observed in PTSD.

Acknowledgments

References

Chen VM, Foilb AR, Christianson JP. Inactivation of ventral hippocampus interfered with cued-fear acquisition but did not influence later recall or discrimination. Behav Brain Res. 2016; 296:249–253. [PubMed: 26386302]

Christianson JP, Jennings JH, Ragole T, Flyer JG, Benison AM, Barth DS, Watkins LR, Maier SF. Safety signals mitigate the consequences of uncontrollable stress via a circuit involving the sensory insular cortex and bed nucleus of the stria terminalis. Biol Psychiatry. 2011; 70:458–464. [PubMed: 21684526]

Clarke HF, Horst NK, Roberts AC. Regional inactivations of primate ventral prefrontal cortex reveal two distinct mechanisms underlying negative bias in decision making. Proc Natl Acad Sci U S A. 2015; 112:4176–4181. [PubMed: 25775597]

Dalley JW, Cardinal RN, Robbins TW. Prefrontal executive and cognitive functions in rodents: neural and neurochemical substrates. Neurosci Biobehav Rev. 2004; 28:771–784. [PubMed: 15555683]

Foilb AR, Bals J, Sarlitto MC, Christianson JP. Sex differences in fear discrimination do not manifest as differences in conditioned inhibition. Learn Mem. 2018; 25:49–53. [PubMed: 29246981]

Foilb AR, Christianson JP. Serotonin 2C receptor antagonist improves fear discrimination and subsequent safety signal recall. Prog Neuropsychopharmacol Biol Psychiatry. 2016; 65:78–84. [PubMed: 26344640]

Foilb AR, Flyer-Adams JG, Maier SF, Christianson JP. Posterior insular cortex is necessary for conditioned inhibition of fear. Neurobiol Learn Mem. 2016; 134(Pt B):317–327. [PubMed: 27523750]

Gallagher M, McMahan RW, Schoenbaum G. Orbitofrontal cortex and representation of incentive value in associative learning. J Neurosci. 1999; 19:6610–6614. [PubMed: 10414988]

Gallistel CR. Animal cognition: the representation of space, time and number. Annu Rev Psychol. 1989; 40:155–189. [PubMed: 2648974]

George P, Charles W. The Rat Brain in Stereotaxic Coordinates. 2013

Gremel CM, Costa RM. Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. Nat Commun. 2013; 4:2264. [PubMed: 23921250]

Grillon C, Morgan CA. Fear-potentiated startle conditioning to explicit and contextual cues in Gulf War veterans with posttraumatic stress disorder. J Abnorm Psychol. 1999; 108:134–142. [PubMed: 10066999]

Hakamata Y, Matsuoka Y, Inagaki M, Nagamine M, Hara E, Imoto S, Murakami K, Kim Y, Uchitomi Y. Structure of orbitofrontal cortex and its longitudinal course in cancer-related post-traumatic stress disorder. Neurosci Res. 2007; 59:383–389. [PubMed: 17923164]

Izquierdo A. Functional Heterogeneity within Rat Orbitofrontal Cortex in Reward Learning and Decision Making. J Neurosci. 2017; 37:10529–10540. [PubMed: 29093055]

Izquierdo A, Suda RK, Murray EA. Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. J Neurosci. 2004; 24:7540–7548. [PubMed: 15329401]

Jason H. Multiple Comparisons. 1996

Jenewein J, Erni J, Moergeli H, Grillon C, Schumacher S, Mueller-Pfeiffer C, Hassanpour K, Seiler A, Wittmann L, Schnyder U, Hasler G. Altered Pain Perception and Fear-Learning Deficits in Subjects With Posttraumatic Stress Disorder. J Pain. 2016; 17:1325–1333. [PubMed: 27641312]

Jovanovic T, Kazama A, Bachevalier J, Davis M. Impaired safety signal learning may be a biomarker of PTSD. Neuropharmacology. 2012; 62:695–704. [PubMed: 21377482]

Jovanovic T, Norrholm SD, Fennell JE, Keyes M, Fiallos AM, Myers KM, Davis M, Duncan EJ. Posttraumatic stress disorder may be associated with impaired fear inhibition: relation to symptom severity. Psychiatry Res. 2009; 167:151–160. [PubMed: 19345420]

Kahnt T, Tobler PN. Salience signals in the right temporoparietal junction facilitate value-based decisions. J Neurosci. 2013; 33:863–869. [PubMed: 23325225]

Liberzon I, Martis B. Neuroimaging studies of emotional responses in PTSD. Ann N Y Acad Sci. 2006; 1071:87–109. [PubMed: 16891565]

Lichtenberg NT, Pennington ZT, Holley SM, Greenfield VY, Cepeda C, Levine MS, Wassum KM. Basolateral Amygdala to Orbitofrontal Cortex Projections Enable Cue-Triggered Reward Expectations. J Neurosci. 2017; 37:8374–8384. [PubMed: 28743727]

Majchrzak M, Di Scala G. GABA and muscimol as reversible inactivation tools in learning and memory. Neural Plast. 2000; 7:19–29. [PubMed: 10709211]

McDannald MA, Saddoris MP, Gallagher M, Holland PC. Lesions of orbitofrontal cortex impair rats' differential outcome expectancy learning but not conditioned stimulus-potentiated feeding. J Neurosci. 2005; 25:4626–4632. [PubMed: 15872110]

Murray EA, O'Doherty JP, Schoenbaum G. What we know and do not know about the functions of the orbitofrontal cortex after 20 years of cross-species studies. J Neurosci. 2007; 27:8166–8169. [PubMed: 17670960]

Ogawa M, van der Meer MA, Esber GR, Cerri DH, Stalnaker TA, Schoenbaum G. Risk-responsive orbitofrontal neurons track acquired salience. Neuron. 2013; 77:251–258. [PubMed: 23352162]

Ongür D, Price JL. The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. Cereb Cortex. 2000; 10:206–219. [PubMed: 10731217]

Ostlund SB, Balleine BW. Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning. J Neurosci. 2007; 27:4819–4825. [PubMed: 17475789]

Ostroff LE, Cain CK, Bedont J, Monfils MH, Ledoux JE. Fear and safety learning differentially affect synapse size and dendritic translation in the lateral amygdala. Proc Natl Acad Sci U S A. 2010; 107:9418–9423. [PubMed: 20439732]

Price JL. Definition of the orbital cortex in relation to specific connections with limbic and visceral structures and other cortical regions. Ann N Y Acad Sci. 2007; 1121:54–71. [PubMed: 17698999]

Rempel-Clower NL. Role of orbitofrontal cortex connections in emotion. Ann N Y Acad Sci. 2007; 1121:72–86. [PubMed: 17846152]

Rodriguez-Romaguera J, Do-Monte FH, Tanimura Y, Quirk GJ, Haber SN. Enhancement of fear extinction with deep brain stimulation: evidence for medial orbitofrontal involvement. Neuropsychopharmacology. 2015; 40(7):1726–33. DOI: 10.1038/npp.2015.20 [PubMed: 25601229]

Rogan MT, Leon KS, Perez DL, Kandel ER. Distinct neural signatures for safety and danger in the amygdala and striatum of the mouse. Neuron. 2005; 46:309–320. [PubMed: 15848808]

Rudebeck PH, Murray EA. The orbitofrontal oracle: cortical mechanisms for the prediction and evaluation of specific behavioral outcomes. Neuron. 2014; 84:1143–1156. [PubMed: 25521376]

Sangha S, Chadick JZ, Janak PH. Safety encoding in the basal amygdala. J Neurosci. 2013; 33:3744–3751. [PubMed: 23447586]

Sangha S, Robinson PD, Greba Q, Davies DA, Howland JG. Alterations in reward, fear and safety cue discrimination after inactivation of the rat prelimbic and infralimbic cortices. Neuropsychopharmacology. 2014; 39:2405–2413. [PubMed: 24727732]

Schoenbaum G, Nugent SL, Saddoris MP, Setlow B. Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. Neuroreport. 2002; 13:885–890. [PubMed: 11997707]

Shiba Y, Santangelo AM, Roberts AC. Beyond the Medial Regions of Prefrontal Cortex in the Regulation of Fear and Anxiety. Front Syst Neurosci. 2016; 10:12.doi: 10.3389/fnsys.2016.00012 [PubMed: 26941618]

Stalnaker TA, Cooch NK, Schoenbaum G. What the orbitofrontal cortex does not do. Nat Neurosci. 2015; 18:620–627. [PubMed: 25919962]

Trow JE, Hong NS, Jones AM, Lapointe J, MacPhail JK, McDonald RJ. Evidence of a role for orbital prefrontal cortex in preventing over-generalization to moderate predictors of biologically significant events. Neuroscience. 2017; 345:49–63. [PubMed: 27746344]

Walton ME, Behrens TE, Buckley MJ, Rudebeck PH, Rushworth MF. Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. Neuron. 2010; 65:927–939. [PubMed: 20346766]

Wilson RC, Takahashi YK, Schoenbaum G, Niv Y. Orbitofrontal cortex as a cognitive map of task space. Neuron. 2014; 81:267–279. [PubMed: 24462094]

Wright KM, DiLeo A, McDannald MA. Early adversity disrupts the adult use of aversive prediction errors to reduce fear in uncertainty. Front Behav Neurosci. 2015; 9:227. [PubMed: 26379520]

Yehuda R, Hoge CW, McFarlane AC, Vermetten E, Lanius RA, Nievergelt CM, Hobfoll SE, Koenen KC, Neylan TC, Hyman SE. Post-traumatic stress disorder. Nat Rev Dis Primers. 2015; 1:15057. [PubMed: 27189040]

Zelinski EL, Hong NS, Tyndall AV, Halsall B, McDonald RJ. Prefrontal cortical contributions during discriminative fear conditioning, extinction, and spontaneous recovery in rats. Exp Brain Res. 2010; 203:285–297. [PubMed: 20449729]

Zimmermann KS, Li CC, Rainnie DG, Ressler KJ, Gourley SL. Memory Retention Involves the Ventrolateral Orbitofrontal Cortex: Comparison with the Basolateral Amygdala. Neuropsychopharmacology. 2018; 43:373–383. [PubMed: 28664926]

Highlights

- The neural mechanisms that allow for differential responses to alternating danger and safety signals are unknown.

- Inactivation of ventrolateral orbitofrontal cortex impaired fear inhibition during fear discrimination.

- Ventrolateral orbitofrontal cortex is critical to fear inhibition by safety signals.
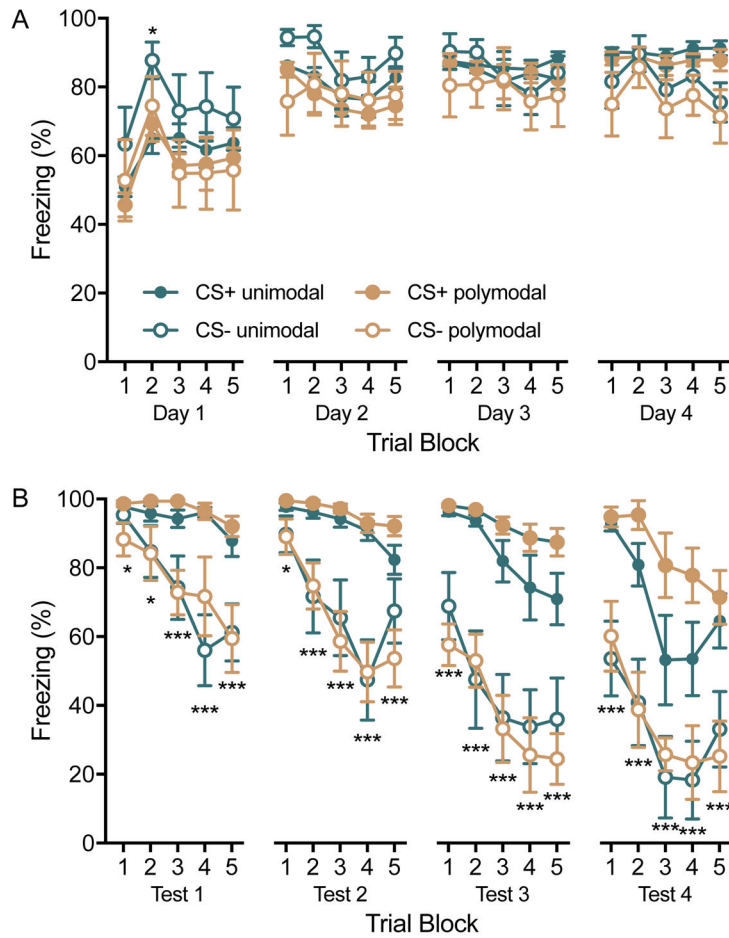
Figure 1.
Freezing to conditioned stimuli (CS) during daily conditioning sessions and recall tests. Rats ($n = 8$/group) were assigned to conditioning with auditory stimuli (unimodal) or auditory and visual stimuli (polymodal). Each day of conditioning entailed 15 CS+ and 15 CS− presentations and occurred in the afternoon. Recall tests entailed 30s presentations of the CSs with a 30s inter-trial-interval and were given each morning, beginning on Day 2. Although discrimination was not evident during conditioning sessions, rats demonstrated behavioral flexibility in response to alternating CS+ and CS− trials in which freezing levels modulated with the expectation of shock. Cue modality did not significantly influence freezing during conditioning or recall. **(A)** Mean(+/−S.E.M.) freezing in blocks of 3 CS presentations during conditioning sessions. *$p < 0.05$ Trial 2 compared to trials 1 and 3. **(B)** Mean(+/−S.E.M.) freezing in blocks of 2 CS presentations during recall tests. *$p < 0.05$, ***$p < 0.001$ freezing to CS− significantly less than CS+.
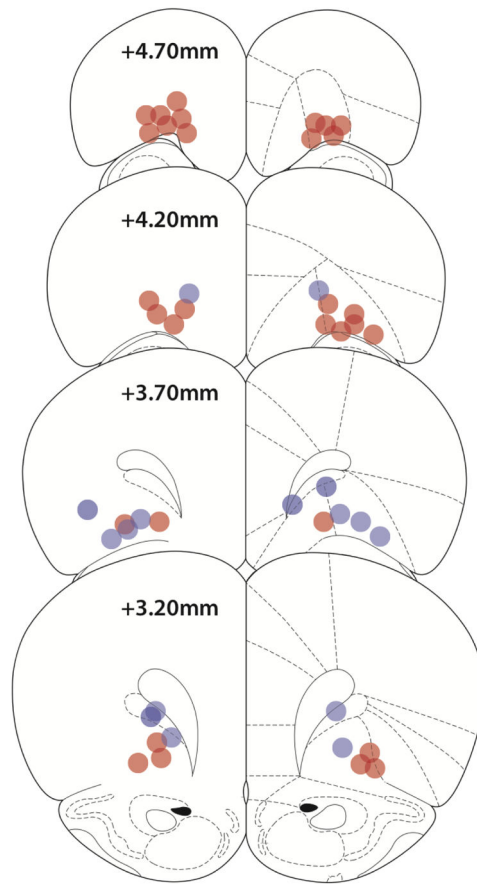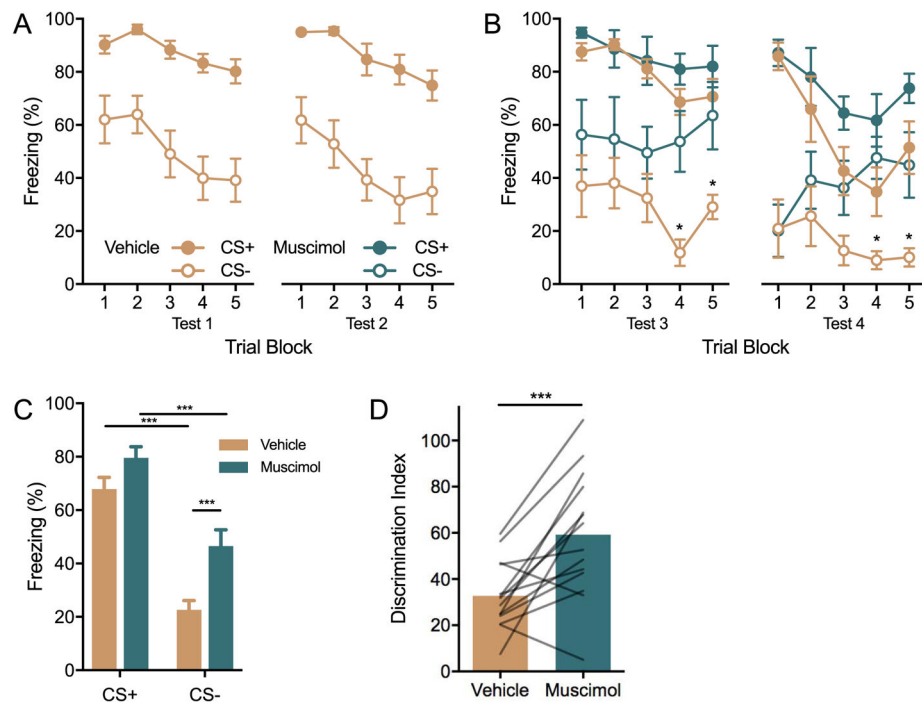
Figure 2.
Microinjection cannula tip locations of rats included in Experiment 2 (blue circles) and
Experiment 3 (red circles).
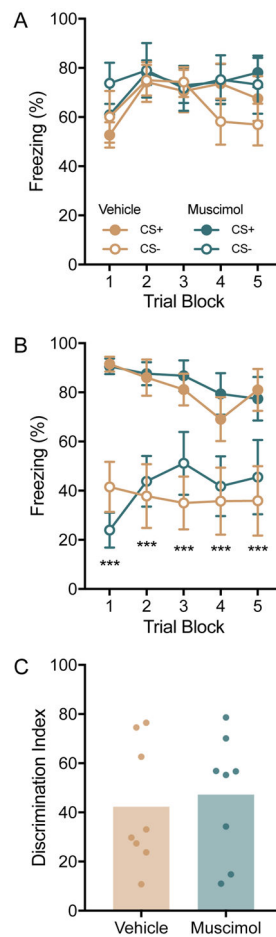
Figure 3.
Behavioral freezing in Experiment 2. Rats ($n = 8$/group) were assigned to either intra-OFC muscimol or vehicle injections which were made 60 min before fear discrimination conditioning. **(A)** Mean(+/− S.E.M.) freezing to the CSs during blocks of 3 conditioning trials. Pretreatment with muscimiol did not alter freezing compared to vehicle controls. **(B)** Mean(+/−S.E.M.) freezing to the CSs (2 trials per block) in a recall test given 24h after discrimination conditioning. There was no effect of pre-conditioning muscimol on later discrimination in the recall test. ***$p < 0.001$ freezing to the CS− was significantly less than the CS+ across all trial blocks. **(C)** Mean (individual replicates) discrimination index (Freezing to the CS− divided by the CS+ times 100). Muscimol and vehicle treated rats behaved equivocally in the discrimination task.

Figure 4.

Inactivation of the OFC interfered with recall of conditioned fear discrimination. **(A)** Mean(+/−S.E.M) freezing to the CS+ and CS− in blocks of 2 discrimination recall trials in tests 1 and 2 prior to drug administration. Robust discrimination was evident with significantly reduced freezing to the CS− at every trial, *ps < 0.004.* Legend identifies treatments in both Fig. 4A and 4B; in Fig 4A "vehicle" treatment indicates behavior prior to any drug administration. **(B)** Mean(+/−S.E.M.) freezing to the CS+ and CS− in blocks of 2 trials 60 min after injection of muscimol or vehicle (*n*s = 7/group/test). Discrimination was evident in both tests (see text for significant contrasts), but greater discrimination was evident in the vehicle groups in the later trials. *$ps < 0.019$ CS− vehicle vs. CS− muscimol. **(C)** Mean(+S.E.M.) freezing to the CS+ and CS− pooled across test days 3 and 4, and trials. Differential freezing was significant in both vehicle and muscimol conditions but there was greater freezing expressed to the CS− in muscimol treated animals. ***$p < 0.001$ differences between means as indicated by overhead lines. **(D)** Mean fear discrimination index (lines indicate individual replicates). ***$p < 0.001$ suppression of freezing to the CS− was significantly better in the vehicle group.