

# Medial orbitofrontal cortex modulates associative learning between environmental cues and reward probability

**Authors:** Sam Hall-McMaster, Jessica Millar, Ming Ruan, Ryan Ward

This work originated from The University of Otago. PO Box 56, Dunedin 9054, New Zealand

**Corresponding author:** Ryan Ward. Department of Psychology, University of Otago  
PO Box 56, Dunedin 9054, New Zealand. rward@psy.otago.ac.nz

**Running title:** mOFC and reward probability learning

**Number of pages:** 37

**Number of figures:** 7 (in manuscript); 2 (in supplementary materials)

**Number of tables:** 1 (in supplementary materials)

**Word count (introduction-discussion, inclusive):** 6669

**Word count (abstract):** 199

**Competing interests:** The authors declare no conflict of interest.

**Keywords:** rat medial orbitofrontal cortex; neuronal excitation; reward probability learning; sustained attention; signaled reward probability

### **Abstract**

It has recently been recognised that orbitofrontal cortex has two subdivisions that are anatomically and functionally distinct. Most rodent research has focused on the lateral subdivision, leaving the medial subdivision (mOFC) relatively unexplored. We recently showed that inhibiting mOFC neurons eliminated the differential impact of reward probability cues on discrimination accuracy in a sustained attention task. In the present study, we tested whether increasing mOFC neuronal activity in rats would accelerate acquisition of reward contingencies. mOFC neuronal activity was increased using the DREADD (Designer Receptors Exclusively Activated by Designer Drugs) method, in which clozapine-N-oxide administration leads to neuronal modulation by acting on synthetic receptors not normally expressed in the rat brain. We predicted that rats with neuronal activation in mOFC would require fewer sessions than controls for acquisition of a task in which visual cues signal the probability of reward for correct discrimination performance. Contrary to this prediction, mOFC neuronal activation impaired task acquisition, suggesting mOFC may play a role in learning relationships between environmental cues and reward probability or for using that information in adaptive decision-making. In addition, disrupted mOFC activity may contribute to psychiatric conditions in which learning associations between environmental cues and reward probability is impaired.

## **Introduction**

Orbitofrontal cortex (OFC) function is critical to adaptive decision-making (Rushworth, Noonan, Boorman, Walton, & Behrens, 2011). In particular, evidence suggests that OFC supports reward-guided behaviour by encoding (or updating) the value of predicted outcomes or supporting behavioural adaptation based on these encoded values (Roesch & Olson, 2005; Burton, Kashtelyan, Bryden, & Roesch, 2014).

OFC has two anatomical subdivisions that are thought to be functionally distinct (Rushworth et al., 2011; Zald et al., 2012). Lateral OFC (lOFC) is thought to be involved in assigning value to particular options (Noonan et al., 2010; Rushworth et al., 2011). By contrast, mOFC is thought to make use of learned representations to guide behaviour (Rushworth, Kolling, Sallet, & Mars, 2012) such as focusing attention on relevant task aspects (Walton, Behrens, Noonan, & Rushworth, 2011).

Although the specific role of mOFC is relatively unexplored in rodents, it has been shown that inactivating mOFC in mice shifts choice from small-certain rewards to larger-uncertain rewards (Stopper, Green, & Floresco, 2014) and increases willingness to expend effort to obtain rewards (Gourley, Lee, Howell, Pittenger, & Taylor, 2010). In addition, neurons in mOFC have been shown to increase firing in response to cues that predict low value outcomes (Burton et al., 2014). Thus, current research in rodents suggests mOFC plays a role in governing adaptive behaviour in situations with dynamic relationships between behaviour and rewarding outcomes.

Most research on OFC function has focused on how encoded value impacts relatively simple responses and few studies have addressed the role of OFC in the ability of reward-associated cues to modulate higher order cognitive processes. One paradigm that seeks to bridge this research gap is the signaled probability sustained

attention task (SPSA; Ward et al., 2015a,b), which is modeled after the five-choice serial reaction time task (Robbins, 2002). In the SPSA, animals are situated in operant boxes where they must press a previously cued lever to obtain reward (see Figure 1). On each trial, the reward probability for pressing the previously cued lever is signaled as being high or low (probability = 1.0 or 0.1 respectively). The reward probability on a given trial is signaled by the state of the houselight (on or off). Attentional load is manipulated on the task by varying cue light duration so that maintaining attention over the entire session requires effort (Ward et al., 2015a,b). Thus optimal performance on this task is achieved through learning a behavioural strategy that leads to increased accuracy on high reward probability trials (in which rewards are guaranteed for a correct response) relative to low reward probability trials.

Using this paradigm, we recently showed that preferentially inhibiting neural activity in mOFC attenuated the impact of reward-associated cues on response accuracy in animals that had been trained on the SPSA (Ward et al., 2015b). More specifically, preferential inhibition of mOFC neurons abolished animals' ability to use reward probability cues to show increased accuracy on high relative to low reward probability trials. This did not occur by affecting the ability of animals to encode the relationship between predictive cues and outcomes, as evidenced by an ability to discriminate between trial types based on choice response latencies and omissions (Ward et al., 2015b). These data suggest that inhibiting mOFC neuronal activity selectively impaired the ability of animals to use predictive cues to guide adaptive behaviour. If this were the case, it is possible that activating mOFC neurons may enhance this ability and in turn accelerate learning of reward probability contingencies.

The present experiment aimed to test this possibility by investigating whether neuronal activation in mOFC would accelerate task acquisition on the SPSA (Ward et al., 2015a,b). We predicted that rats with neuronal activation of mOFC neurons would reach an acquisition criterion in fewer sessions than control animals. This prediction was based on several lines of evidence. Firstly, manipulating mOFC activity has been shown to increase effort to obtain reward (Gourley et al., 2010). Maximising reward in the SPSA requires learning the contingencies governing the relationship between cues and the probability of receiving a reward, a process thought to be mediated by attention (Ward et al., 2015a,b). Thus increased attentional effort on high reward probability trials could contribute to faster acquisition. Secondly, reducing mOFC activity has been shown to increase risky decision-making (Stopper et al., 2014). Increasing mOFC neuronal activity may therefore inhibit the selection of risky options, allowing animals to maintain low risk but high payoff choices. In the context of SPSA, this would be reflected as increased choice accuracy on high payoff trials relative to low payoff trials, an outcome that forms the basis of acquisition. Thirdly, it has previously been shown that mOFC neurons fire more for cues that predict less rewarding options (Burton et al., 2014). Increasing mOFC neuronal activity may therefore boost the salience of less rewarding environmental cues, improve discrimination between trial types and accelerate SPSA acquisition. All of these processes contribute to the acquisition of our task and thus, the manipulation of any of these psychological processes via activation of mOFC neurons could accelerate learning within the SPSA paradigm.

## **Materials and Methods**

### **Subjects**

Subjects were 20 male Long-Evans rats (Hercus Taieri Resource Unit, Dunedin, New Zealand), housed one to three per cage in a temperature-controlled room ( $22^{\circ}\text{C} \pm 1^{\circ}\text{C}$ ). Rats were kept on a 12-hour light-dark cycle, with water available ad libitum. Animals were aged between 70 and 88 days when surgeries began and all experimental procedures were approved by the University of Otago Animal Ethics Committee.

### **Apparatus**

The experiment was conducted using 10 identical operant chambers (Med-associates, St. Albans, VT: model ENV-008w) with internal dimensions of 30.5cm (length) x 24.1cm (width) x 21.0cm (height). One wall of each chamber consisted of cue lights located above two retractable levers and a food hopper centered between them. The opposite wall contained a houselight, which provided chamber illumination, and a speaker, which presented tones to signal reward delivery (90dB, 2500Hz, 200ms). Head entries into the food hopper were recorded using an infrared photocell detector. Ambient noise was attenuated by a fan, which produced 72dB white noise. Chambers were situated in individual cupboards to block ambient light and restrict external noise. Experimental events were programmed and data were recorded using MedPC IV.

### **Procedure**

**Surgery.** mOFC coordinates were determined using the rat brain atlas (Paxinos & Watson, 2007) in consultation with previous studies targeting mOFC in rats (Fuchs, Evans, Parker, & See, 2004; Malkuz et al., 2015), especially in male Long-Evans (Mar, Walker, Theobald, Eagle, & Robbins, 2011; Burton et al., 2014,

Stopper et al., 2014). Rats were anaesthetised with Ketamine (75mg/kg), Domitor (0.5mg/kg) and given 1µL stereotaxic, bilateral virus injections into the mOFC at coordinates: +4.00mm relative to bregma; ±0.6mm lateral to the midline; and -4.25mm below the brain surface. AAV2/hSyn-HA-hM3D(Gq)-IRES-mCitrine viruses (hereafter referred to as hM3D(Gq);  $2 \times 10^{12}$  particles/ml) were obtained from the Gene Therapy Center Vector Core (University of North Carolina) and were injected using a Hamilton syringe attached to a syringe pump at an infusion rate of 0.25µl/min. All subjects received hM3D(Gq) viral injections because studies using DREADD (including our own previous studies) have repeatedly shown that viral injection alone, and DREADD expression in the absence of CNO, do not affect cellular activity (Armbruster, Li, Pausch, Herlitze, & Roth, 2007; Pei, Rogan, Yan, & Roth, 2008; Nichol & Roth, 2009) or behaviour (Krashes et al., 2011; Parnaudeau et al., 2013, 2015; Ward et al., 2015b).

Following surgery, rats were given Antisedan (2.5mg/kg) to reverse the anaesthetic effects of Domitor, as well as Amphoprim (0.2ml) and Caprieve (5mg/kg) subcutaneously for pain relief. Rats were individually housed and given oral Caprieve for three days, as required. Rats were then returned to their regular housing and given a minimum of 7 days recovery time before beginning food deprivation. Training began after rats reached 85-90% free feeding weight and were kept in this weight range by weighing daily and adjusting food intake accordingly.

**Sustained-attention task training.** First, rats were accustomed to pellets to be used as a reward during the experiment by placing five pellets per rat per cage for two days before training onset. Rats were trained during the light phase, once per day, seven days per week, based on steps outlined by Ward and colleagues (2015a,b). Sessions were conducted at approximately the same time daily. Rats were first trained

to consume pellets from the food hopper. Rats received two sessions in which 60 pellets were delivered on a variable interval schedule (mean duration = 30s, range = 0.76s - 119.87s).

**Lever-press training.** Rats were then trained to press levers to obtain reward under continuous reinforcement (CRF). In these trials, the left or right lever was extended for 10s, after which the lever was retracted and a pellet was delivered. Pressing the lever within 10s caused the lever to retract, triggered pellet delivery into the hopper and initiated an inter-trial interval (ITI; mean duration = 30s, range = 0.76s - 119.87s) in which the houselight was turned off. Each session consisted of 30 left and 30 right lever extensions, presented pseudorandomly. This ensured no more than four consecutive presentations of the same trial type and was applied to every training stage hereafter. Rats completed three CRF sessions. On the third session all subjects made more than 50 presses out of 60 trials.

**Single cue-single lever training.** In these trials, the cue light above either the left or right lever was illuminated for 5s. The cue was terminated and 1s later, the lever beneath the cue light was extended for 10s. Pressing the extended lever caused the lever to retract, triggered reward delivery and initiated a new trial, whereas failure to press (an omission) directly initiated a new trial. From this stage onwards the variable time between cue presentations (hereafter called a pre-cue interval) was increased to 45s (range = 2.74s – 148.13) and sessions lasted for 68 trials. Rats had three sessions on this training stage. The first and second sessions presented only the left or right cue light/lever. The third session consisted of 50% left and 50% right cue light/lever trials. All rats pressed the extended lever more than 65/68 times on the third session.



**Choice training.** In choice trials, one lever was cued for 5s and, following a 1s delay, both levers were presented for 10s. A response on either lever resulted in retraction of both levers and responses to the previously cued lever were rewarded. Choice trials were intermingled with single-cue trials described above. Rats were trained for three sessions with 50% choice trials, one session with 75% and three sessions with 100%. Incorrect responses resulted in a correction procedure, in which trials were repeated, with the same lever being cued, until a correct response was made. Rats then had three sessions of 100% choice trials during which the correction procedure was turned off. Under these conditions, an incorrect response or omission, resulted in lever retraction and initiation of the next trial.

**Decreasing cue duration.** For the purposes of the present experiment, it was desirable to ensure the task was difficult enough to recruit attentional effort and avoid ceiling effects in performance during the testing stage. Given cue duration modulates accuracy on this task (Ward et al., 2015a,b) we reduced cue duration from 5s to 1s over 8 sessions (4 sessions at 2s and 4 sessions at 1s). Before each session in this training stage, rats were familiarised with the injection procedure by holding them in the intraperitoneal (i.p.) injection position and gently probing the abdomen with a needleless syringe. Rats that failed to score above 80% on the fourth 1s cue session were given a fifth session. This included three Saline rats and two CNO+ rats (see results for analysis-driven CNO+/CNO- distinctions).

**SPSA.** In the final version of the task (Figure 1), each trial began with the houselight in a particular state (on or off) in which it remained for the duration of the trial (until a choice response was made). The state of the houselight signaled reward probability for making a correct response (1.0 or 0.1; counterbalanced across rats). The rest of the task followed the previous training stage. Following a pre-cue interval,

one of the levers was cued for 1s and, following a 1s delay, both levers were presented for a maximum of 10s. Pressing one lever within this time frame caused both levers to retract. Rats were rewarded for pressing the previously cued lever, whereas an incorrect response or omission resulted in no reward. Equal numbers of high and low probability trials were presented in each session.

Following a between-subjects design, rats were divided into two groups, which were counterbalanced based on accuracy scores from the final training session. In cases where animals received an additional session for failing to meet the 80% training criteria, accuracy scores from the additional training session were used for counterbalancing. Each group was randomly assigned to saline or CNO conditions. In the presence of hM3D(Gq) receptor, CNO administration has been shown to depolarise neurons in vitro (Alexander et al., 2009; Krashes et al., 2011; Kong et al., 2012; Nakajima et al., 2016) and increase neuronal activity both in vitro (Krashes et al., 2011; Kong et al., 2012) and in vivo (Alexander et al., 2009; Kong et al., 2012). Importantly, the CNOactivated hM3D(Gq) DREADD receptor has been shown to significantly increase activity in prefrontal cortex neurons (Carreno et al., 2015; Yau & McNally, 2015) and has previously been used to interrogate mOFC function with regard to goal directed behaviour (Gourley, Zimmermann, Allen, & Taylor, 2016). In addition, a large and growing body of evidence demonstrates that CNO administered to rodents without DREADD expression does not impact behaviour (Roth, 2016; Parnaudeau et al., 2013; Bradfield, Dezfouli, van Holstein, Chieng, & Balleine, 2015; Carreno et al., 2015; Parnaudeau et al., 2015; Yau & McNally, 2015; Gourley et al., 2016), including SPSA performance (Ward et al., 2015b). Rats received daily i.p. injections of saline or CNO based on doses used previously (2.0mg/kg; Ward et al., 2015b, Parnaudeau et al., 2013, 2015; volume = 0.5ml/kg). CNO was prepared fresh

daily. Rats were injected 30 mins before being tested on the SPSA (e.g. Ward et al., 2015a,b) and performed one session daily for a total of 20 sessions. Our previous results (Ward et al., 2015a, b) along with pilot work indicated that this number of sessions was sufficient to assess the acquisition of the task.

**Histology.** Following the conclusion of the experiment, rats were deeply anaesthetised and perfused with paraformaldehyde and brains processed for histology to verify injection location and spread of viral infection (see Figure 2).

### **Data Analysis**

The main dependent variable of interest was proportion correct. During the SPSA phase, acquisition was defined as proportion correct being significantly greater on high compared with low reward probability trials over three consecutive sessions, as assessed by paired t-tests. In most cases, the acquisition criterion was met from the first three sessions in which accuracy on high probability trials was greater than on low probability trials. For the cases where multiple comparisons were conducted, a Bonferroni correction was employed. Acquisition distributions were compared using a Kolmogorov-Smirnov test. Latency to make a choice response, the number of omissions on high and low probability trials, as well as the number of perseverative responses and errors were also measured. Where appropriate, data were analysed by repeated measures ANOVA and group means were compared using unpaired t-tests.

## **Results**

### **No group differences were observed prior to the experimental manipulation**

Bilateral stereotaxic injection of hM3D(G<sub>q</sub>)-mCitrine expressing adeno-associated viruses resulted in hM3D(G<sub>q</sub>) and mCitrine expression, which were made selective to neurons through the use of the human synapsin1 promotor (hSyn). Figure 2a shows a representative image of viral expression in the mOFC. Figure 2b shows

the minimal (black) and maximal (grey) extent of intrinsic fluorescence of mCitrine (from the hM3D(Gq)-expressing virus). Viral expression was focused on mOFC, with spreading in some cases to ventral OFC and no expression in lateral OFC. Two rats from the Saline group were excluded from subsequent analyses based on lack of expression, or expression in areas outside of medial orbitofrontal cortex. The pattern of viral expression did not differ systematically across rats that received saline or CNO in the subsequent testing phase, either between or within groups (data from CNO rats that did and did not acquire the task, CNO+ and CNO-, respectively, are shown in Figure S2).

Figure 3a shows the proportion correct (accuracy) over the choice and decreasing cue duration stages of training. Data are presented separately for groups of rats that would receive either saline or CNO in the testing phase. For this analysis, data were further subdivided according to CNO rats that did and did not acquire in the experimental phase (CNO+ and CNO-, respectively). For all groups, accuracy rose from approximately 80% over the first five sessions to approximately 90% when 100% choice trials were reached. When cue duration was decreased from 5s to 1s, over eight sessions, accuracy was reduced to approximately 80%, indicating this manipulation was effective in producing a greater attentional load. In general, performance between groups was not different across training phases. To verify these visual impressions of the figure, we averaged scores within each training sub-stage separately for each group. A 3 x 6 repeated measures ANOVA on these data, with group as a between-subjects factor and training stage as a within-subjects factor, revealed a main effect of training stage ( $F(5, 75) = 15.676, p = 0.00$ ) but no main effect of group ( $F(2, 15) = 0.202, p = 0.819$ ) and no Training Stage x Group Interaction ( $F(10, 75) = 1.016, p = 0.438$ ).

Figure 3b shows accuracy on the final session of choice training for all groups of animals. As above, data are shown separately for CNO rats that did and did not acquire in the experimental phase. In cases where animals received an additional session for failing to meet the 80% training criteria, accuracy in the additional session is used as that animal's contribution to the mean accuracy for a given group. All three groups of rats scored on average between approximately 80 and 85% correct on the final training session, with CNO- rats performing somewhat better than Saline and CNO+ rats. We conducted a one-way ANOVA on the accuracy scores from the three groups. This analysis found no significant difference between the groups ( $F(2,15)=2.65$ ,  $p=0.26$ ). To discount the possibility that the number of rats housed together influenced the results, we conducted a one-way ANOVA on proportion correct for the last day of choice training with housing number as a between-subjects factor. Differences in performance due to housing number were not significant ( $F(2, 15) = 0.80$ ,  $p = 0.50$ ).

Together, these results indicate all animals acquired the training steps to an equal level of performance, manipulating cue duration was effective in reducing performance to the desired range and groups did not differ in performance prior to the experimental manipulation.

### **mOFC neuronal activation impaired SPSA acquisition**

Figure 4 shows the proportion of animals that acquired the SPSA in CNO and Saline groups over 20 testing sessions. Data are presented using a cumulative distribution plot, which shows the cumulative proportion of rats that met the acquisition criterion as a function of experimental sessions. This plot has the advantage of presenting all of the data from each subject, rather than presenting an

average value for all subjects. Thus, overall performance can be assessed, but individual differences are also readily observed.

Figure 4 indicates the final proportion of animals acquiring the task was greater for the Saline group (1.0; 8 of 8 rats) than for the CNO group (0.5; 5 of 10 rats). In addition, the acquisition distribution for the CNO rats is shifted to the right, indicating the animals in this group that did acquire, did so more slowly than Saline counterparts. To statistically analyse acquisition, we performed a Kolmogorov-Smirnov test, which tests the maximum distance between two distributions and indicates the probability that the distributions come from the same sample population (Massey, 1951). This nonparametric test is useful because there are no assumptions about how the data are distributed (Goodman, 1954). The Kolmogorov-Smirnov analysis revealed a significant difference between the two acquisition distributions ( $D = 0.60, p = 0.00$ ), indicating neuronal activation impaired SPSA acquisition.

### **Rats that failed to acquire could not discriminate between high and low reward probability trials**

Average discrimination accuracy did not differ between any of the groups (see Supplemental Materials). Figure 5 shows proportion correct on high and low probability trials for Saline and CNO groups. Data from the CNO group are separated for rats that did (CNO+) and did not (CNO-) acquire. Data from rats that acquired (Saline/CNO+) were averaged across the three sessions that met the acquisition criterion, whereas data from rats that did not acquire (CNO-) were averaged over the last three test sessions. Saline and CNO+ animals showed greater accuracy on high relative to low probability trials. By contrast, CNO- rats did not show differential accuracy on the two trial types. These results were confirmed statistically by performing a 2 (reward probability) x 3 (group; Saline, CNO+, CNO-) repeated

measures ANOVA on the accuracy data. This analysis revealed a significant effect of reward probability ( $F(1,15) = 70.09, p = 0.00$ ). The effect of group was not significant ( $F(2,15) = 0.132, p = 0.88$ ), but there was a significant reward probability x group interaction ( $F(2,15) = 17.07, p = 0.00$ ).

To locate the source of the interaction, we conducted a separate ANOVA on the data from the two groups that had acquired. This analysis revealed a main effect of reward probability ( $F(1, 11) = 99.09, p = 0.00$ ), however, no main effect of group ( $F(1, 11) = 0.166, p = 0.69$ ), and no interaction ( $F(1,11) = 0.472, p = 0.51$ ). Planned comparisons showed a significant difference between accuracy on high and low reward probability trials for both Saline ( $t(7) = 9.38, p < 0.00$ ) and CNO+ ( $t(4) = 5.30, p = 0.01$ ) rats. By contrast, CNO- rats showed no significant difference in accuracy between trial types ( $t(4) = 0.13, p = 0.90$ ). Additional analyses demonstrated that group differences were not due to a differential effect of houselight state on CNO+ and CNO- animals (see Supplemental Materials) or to changes in perseverative responses or errors (Figure S1).

To further assess the ability of rats that did and did not acquire to discriminate between trial types, we conducted a trial-by-trial analysis of the data to determine whether performance on the current trial was governed by the reward probability on the current or previous trial. Figure 6 shows the results of the analysis for all groups of rats. For those animals that acquired (Saline/CNO+), discrimination accuracy reflects the reward probability on the current trial, regardless of the reward probability on the previous trial; that is, accuracy on high probability trials is always higher than on low probability trials, regardless of the reward probability on the previous trial. By contrast, for rats that did not acquire, discrimination accuracy was non-differential across all trial types. These impressions were confirmed by a 2 (current trial reward

probability) x 2 (previous trial reward probability) x 3 (group) repeated-measures ANOVA which found no main effect of current or previous trial type ( $F_s < 1.5$ ), but a significant interaction of the two ( $F(1,15)=59.24$ ,  $p=0.00$ ). The ANOVA also found a significant interaction between trial type and group ( $F(2,15)=9.16$ ,  $p=0.00$ ). No other main effects or interactions were significant ( $F_s < 1.5$ ). To determine the source of the interaction, we conducted separate ANOVAs on the data from rats that acquired (Saline/CNO+) and those that did not (CNO-). For Saline/CNO+ rats, the ANOVA once again found only a significant interaction between trial type ( $F(1,11)=75.10$ ,  $p=0.00$ ); all other  $p_s > .08$ ), reflecting the prominence of the current trial reward probability signal in governing discrimination accuracy (i.e., accuracy was high on high probability trials and low on low probability trials regardless of the reward probability on the previous trial). For CNO- rats, the ANOVA found no significant main effects or interactions (all  $F_s < 1.0$ ).

In addition to the trial-by-trial analysis above, we examined latency to make a choice response and the number of omissions on high and low probability trials (Table S1). The analysis found no difference in choice response latencies on high or low reward probability trials in any of the groups ( $t_s < 0.50$ ). The analysis of omissions showed all rats completed the vast majority of trials ( $> 90\%$ ). However, Saline/CNO+ rats made significantly more omissions on low compared with high probability trials ( $t(12) = 2.32$ ,  $p = 0.03$ ). By contrast, CNO- rats showed no significant difference in omission number between trial types ( $t(4) = 1.60$ ,  $p = 0.19$ ).

### **Activating mOFC neurons changed the cognitive mechanism of acquisition**

Thus far these analyses have shown that 1) activating mOFC neurons produces a deficit in acquiring our task (only half of CNO rats met acquisition criteria), 2) CNO- rats could not discriminate between high and low reward probability trials, and



3) Saline and CNO+ rats did not differ in their overall performance at the time of acquisition. Thus, for rats that acquired, there were no quantitative differences in performance. Given the previous research cited above on the specific role of the mOFC in modulating behaviour based on the value of reward-associated cues, we reasoned that there may, however, be qualitative differences in learning between Saline and CNO+ rats.

To assess this, Figure 7 shows accuracy on high and low reward-probability trials across the 20 sessions of acquisition (organised in 4-session blocks). As above, the data for the CNO rats are separated according to those that met, or did not meet, acquisition criteria. First, it is important to note that discrimination accuracy did not differ between Saline/CNO+ and CNO- rats in the initial block of training. Proportion correct was between 0.70 and 0.80 for all rats. A 2 (reward probability) x 3 (group) ANOVA conducted on these data found no significant main effects or interactions (all  $p>0.20$ ). Thus, there were no significant differences in performance between groups at the beginning of the acquisition phase.

Figure 7a shows the data from rats that acquired (Saline/CNO+). Proportion correct during the first block was undifferentiated across trial types, and accuracy on high and low probability trials diverged across subsequent session blocks. However, the pattern of divergence was markedly different between Saline and CNO+ rats. For Saline rats, accuracy on low probability trials remained relatively unchanged across session blocks, while accuracy on high probability trials increased. The opposite pattern is seen for CNO+ rats; accuracy on high probability trials remained relatively unchanged while accuracy on low probability trials decreased. For CNO- rats (Figure 7b) accuracy remained between 0.70 and 0.80 across all session blocks, and was undifferentiated across high and low probability trials.

We conducted several analyses to verify the impressions from the figure. First, a reward probability x injection x block ANOVA conducted on the data from Figure 6a found a significant effect of reward probability ( $F(1,11) = 26.02, p = 0.00$ ), and significant reward probability x block ( $F(4,44) = 8.61, p = 0.00$ ) and block x injection ( $F(4,44) = 4.57, p = 0.00$ ) interactions. No other main effects or interactions were significant. Separate ANOVAs on the data from the Saline and CNO+ rats each found significant reward probability x block interactions ( $F(4,28) = 6.92, p = 0.00$ ) and ( $F(4,16) = 3.84, p = 0.02$ ), respectively. These results confirm the divergence of accuracy between high and low probability trials seen in the figure. Second, linear regression analyses of the data from Figure 7a indicated the slope of the best-fitting line for accuracy on high probability trials for Saline rats was positive and significantly different from zero ( $F(1,38) = 14.97, p = 0.00$ ), while the slope of the best fitting line for accuracy on low probability trials for CNO+ rats was negative and significantly different from zero ( $F(1,23) = 6.69, p = 0.02$ ). The slopes of the best fitting lines for accuracy on low probability trials for Saline rats and high probability trials for CNO+ rats were not different from zero ( $F_s < 1.0$ ). Finally, an ANOVA conducted on the data from Figure 7b found no significant effects or interactions ( $F_s < 1.5$ ).

In combination, our analyses indicate that activating mOFC neurons produced a deficit in acquisition of the SPSA task. This deficit was not due to an impact on overall task accuracy or any impact on perseverative responding. In addition, rats meeting acquisition criterion could discriminate between trial types and their discrimination performance was based on the reward probability on the current trial, regardless of injection group. Further support for this result was provided by the finding that animals acquiring the task omitted responding significantly more on low

than on high reward probability trials. Conversely, rats with mOFC neuronal activation that failed to acquire did not discriminate between trial types, as indicated by lack of differential accuracy on high and low probability trials and no difference in the number of omissions between trial types. Finally, Saline rats that acquired the task did so by increasing attention on high probability trials, while CNO+ rats acquired by decreasing attention on low probability trials.

### **Discussion**

Contrary to our original prediction that activating neurons in mOFC would facilitate acquisition of our signaled reward probability task, mOFC neuronal activation impaired SPSA acquisition. This result was initially surprising given mOFC lesions have been previously shown to increase risky decision-making (Stopper et al., 2014). Based on this previous outcome, we reasoned mOFC neuronal activation might decrease risky decision-making, which could be reflected in our task as decreased performance on low probability trials as they bear a large risk of receiving no reward. This was not the case, suggesting both decreases and increases in mOFC activity can hinder adaptive decision-making and that normal mOFC functioning is important for adaptive decisions involving reward probability.

Notwithstanding their surprising nature, our results support and extend previous findings on the role of mOFC in adaptive behaviour. First, our results are consistent with the proposition that increased mOFC activity encodes predictive information about low probability outcomes (Burton et al., 2014). In particular, heightened salience of low value cues, or modulation of attentional resources, produced by activating mOFC neurons may underlie the shift in strategy from increasing accuracy on high probability trials to reducing accuracy on low probability trials shown by CNO+ rats. Second, the present results add to previous findings

showing that reduced mOFC activity in rodents decreases impulsive choice (Mar et al., 2011) and increases goal directed lever pressing (Gourley et al., 2010).

Specifically, our results indicate that, in contrast to the effects of decreased neuronal activity in mOFC, the activation of mOFC neurons can impair adaptive choice. Third, our results extend research by Ward and colleagues (2015b) by suggesting mOFC is not only important for using learned reward probability cue associations to guide behaviour but for learning those associations to begin with. Taken together, too much or too little activity in mOFC neurons seems to impair adaptive decision-making based on information about the likelihood of reward.

### **Mechanism of impaired SPSA acquisition**

One possible psychological mechanism that could underlie the impaired acquisition seen here is the impact of mOFC neuronal activation on attention. We have shown previously that our SPSA task recruits attentional processes, and that increasing attentional load by decreasing cue duration reduces accuracy (Ward et al., 2015a,b). Moreover, the present results indicate that SPSA acquisition under normal conditions involves learning to recruit attentional resources on high reward probability trials, as evidenced by the pattern of acquisition shown by Saline controls. Under conditions of increased neuronal activity in mOFC, the mechanism of acquisition shifted towards reducing accuracy on low reward probability trials, perhaps via learning to reduce attention during these trials. Rats with mOFC neuronal activation were still capable of learning the task, however, this strategy was less adaptive as evidenced by the large number of CNO animals that did not meet the acquisition criterion. One reason this strategy was less adaptive might be that learning the relationship between low probability cues and reward likelihood was more difficult due to fewer trials in which reward was paired with visual cues, relative to

high probability trials. The reason some CNO animals acquired the task and others did not is unclear. There was no systematic variation in viral expression within the CNO group to explain this difference (see Figure S2). Perhaps different acquisition outcomes can be attributed to individual differences, in which some animals were better at compensating for mOFC disruption than others, and with additional training, CNO- animals may have acquired. While it is possible CNO- animals may have acquired given more time, the data in the final block of performance in Figure 7b, provide no evidence that CNO- animals were learning to differentiate performance based on signaled reward probability. Indeed, none of the CNO- rats indicated any improvement of discrimination across the 20 sessions of training. In addition, average accuracy was not reduced by activating neurons in mOFC, arguing against a general deficit in attention, motivation, or perception, and towards a deficit in the ability to modulate attention based on the probability of reward.

It is also evident from the trial-by-trial analysis is that rats acquiring the task were basing their choice behaviour on the reward probability cue from the current trial, regardless of the reward probability cue on the previous trial. This implies that discrimination accuracy was under dynamic control, involving value updating of reward probability on a trial-by-trial basis. By contrast, rats failing to acquire did not differentiate their responses across trial types. In concert with the analysis of perseverative responses and perseverative errors, which showed no differences in perseverative errors after rewarded trials (Figure S1), these data suggest that CNO- rats were likely basing their choice responses on an overall average value associated with the reward-probability signal across trials. The implications of these data for mOFC function are that, under normal conditions, mOFC is engaged in a process of dynamic value updating in response to environmental signals. In some cases,

abnormal mOFC neuronal activity may hamper value updating of reward contingencies and thus impair learning so that the likelihood of reward cannot be distinguished in different contexts.

With further respect to a neural mechanism, the present results could be due to imposing a prolonged state of activity on mOFC neurons (CNO can take 2h to be cleared from the bloodstream in rodents; Guettier et al., 2009). Future research may be able to accelerate acquisition as predicted using a technique with more precise temporal control, such as optogenetics (Gunaydin et al., 2010; Aston-Jones & Deisseroth, 2013). Use of such a technique would also allow us to specify the precise role of mOFC in SPSA acquisition and performance. Specifically, the question of whether mOFC impacts performance during the pre-cue interval (when reward signals are present and attentional resources must be recruited), during the choice phase (when information about signaled reward probability must be combined with attentional information about the location of the correct choice), or after the reward is delivered (when information about the task must be updated), could be greatly informed by such an approach.

In addition, it should be noted that the hSyn promoter used in this study is specific to neurons but not particular neuronal subtypes. Thus, multiple neuronal populations were likely activated upon CNO administration, including both excitatory and inhibitory neurons. Therefore, the present results cannot disentangle the relative contribution of an activation of inhibitory vs excitatory neurons to the behavioural effects seen here. These effects likely reflect an overall disruption of mOFC activity, rather than being due to regional increases in net mOFC activity.

Although viral expression was centered in mOFC, with no systematic differences in expression between groups, it is also conceivable that impaired

acquisition was produced in part by ventral OFC (vOFC) disruption. Given the anterior to posterior anatomy of the OFC subdivisions, some spread to vOFC is common in studies targeting mOFC in rats (see Mar et al., 2011; Burton et al., 2014; Bradfield et al., 2015; Malkuz et al., 2015). Notwithstanding this, mOFC disruption seems a more plausible cause of our results for three reasons. First, viral injections were predominantly located in mOFC, with minor spreading to vOFC. Second, our results are congruent with previous work preferentially manipulating mOFC activity during the SPSA task (Ward et al., 2015b). Third, it has been shown that vOFC/ventrolateral OFC lesions in rats do not impact the acquisition of new discriminations (Brown & Bowman, 2002). Critically, there was no viral spread to IOFC (Figure 2b), a region thought to be involved in associating reward probabilities with certain stimuli (Noonan et al., 2010; Rushworth et al., 2011).

Thus far we have interpreted our results as suggesting that increasing neuronal activity in mOFC impaired learning between environmental cues and reward probability. An alternative interpretation, however, is that mOFC neuronal activation impaired the expression of learned associations at the choice point. For example, artificially increasing neural activity in the mOFC could have masked accurately encoded value signals, such as the relatively low firing rate for high-value cues (Burton et al., 2014) which in turn could prevent the use of these signals to guide adaptive behaviour.

Despite this possibility, the present data argue against a deficit in the expression of learned associations and towards a deficit in the acquisition of reward contingencies. Animals that acquired the task omitted responding significantly more on low probability trials, suggesting these animals had learned to discriminate between trial types using probability cues and responded less on low probability trials

as a result of learning reward was unlikely on these trials. These animals also expressed this learning by making more correct responses on high relative to low probability trials.

If mOFC neuronal activation caused an acquisition deficit, animals that did not acquire would be expected to show no difference in omission number between trial types. This would suggest animals had not learned to discriminate the value of the trial types based on reward probability cues. By contrast, if our results were explained by a deficit in learning expression, animals failing to meet criterion would be expected to omit more low probability trials. This would suggest these animals had still learned to discriminate trial types based on the likelihood of reward. However, the deficit was in using that information to produce greater accuracy on high relative to low probability trials. Our analysis for animals that did not meet acquisition criteria showed no significant difference in omission number between trial types, arguing in favour of the interpretation that activating mOFC neurons disrupted acquisition of the reward probability relationships. Notwithstanding the strength of this evidence, definitive proof of an impact of mOFC neuronal activation on learning vs performance would require a drug free test following the acquisition phase, and future research could take this approach. The present results suggest mOFC may be important for learning to associate environmental cues and reward probability. This interpretation is interesting in terms of the prevailing theoretical view that mOFC plays a relatively minor role in learning to assign value to stimuli, compared with IOFC (Noonan et al., 2010; Rudebeck & Murray, 2014). The present study may challenge this contemporary view of mOFC function, as it provides evidence that mOFC not only compares the value of different choices (Rushworth et al., 2011) but



may be involved in learning the likelihood of reward associated with certain decisions.

### **Relevance to psychiatric disease**

Patients with schizophrenia show deficits in learning to choose options with a higher likelihood of reward (Waltz, Frank, Robinson, & Gold, 2007) and are less likely to choose high-effort, high-reward options than healthy controls, even as reward probabilities increase (Barch, Treadway, & Schoen, 2014). This suggests people with schizophrenia may struggle with learning relationships between their behaviour and reward probability. In a similar manner, our results demonstrated that as a group, rats with activation of mOFC neurons were impaired in learning relationships between reward-associated cues and the likelihood of obtaining reward for their behaviour. A potential implication of these results is that abnormal mOFC activity may contribute to deficits seen in schizophrenia and other conditions characterized by decision-making impairments (e.g. Bechara et al., 2001, Cavedini, Riboldi, Keller, D'Annuncci, & Bellodi, 2002; Sachdev & Malhi, 2005).

### **Acknowledgements**

This research was supported by a University of Otago Research Grant. The authors would like to acknowledge the contributions of Jeremy Anderson (lab manager) and Sara-Lee Illingworth (animal technician) to this investigation. The authors report no conflict of interest.

## References

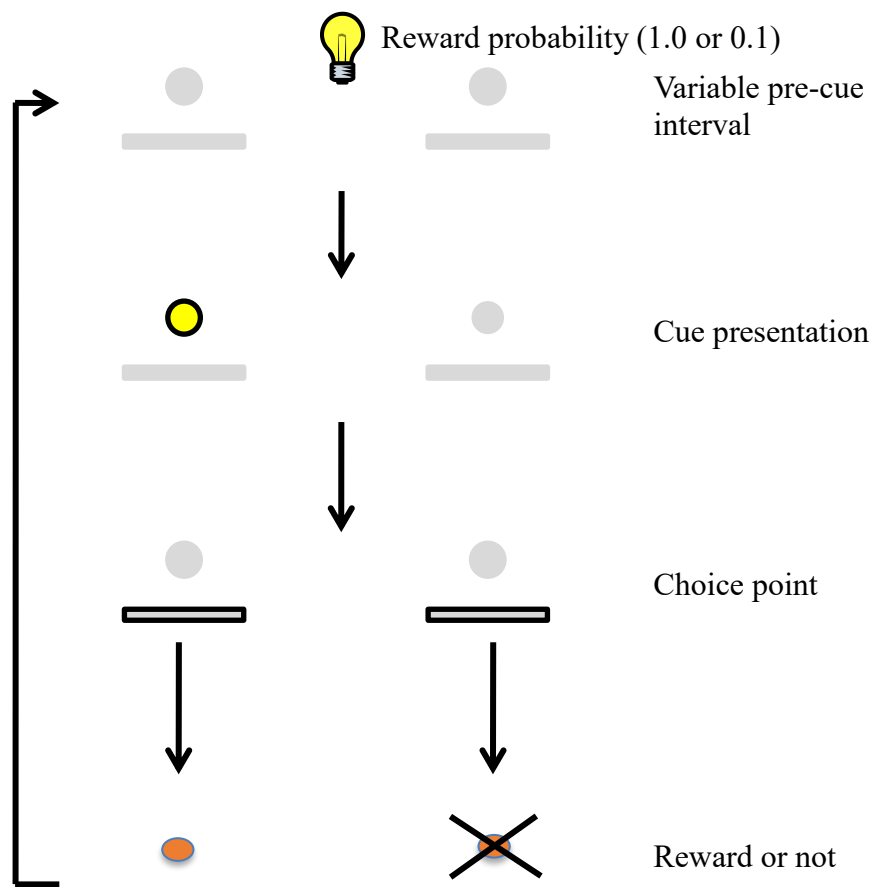
- Alexander, G.M., Rogan, S.C., Abbas, A.I., Armbruster, B.N., Pei, Y., Allen, J.A., ... & McNamara, J.O. (2009) Remote control of neuronal activity in transgenic mice expressing evolved G protein-coupled receptors. *Neuron*, 63, 27-39.
- Armbruster, B.N., Li, X., Pausch, M.H., Herlitze, S. & Roth, B.L. (2007) Evolving the lock to fit the key to create a family of G protein-coupled receptors potentially activated by an inert ligand. *Proc. Nat. Acad. Sci.*, 104, 5163-5168.
- Aston-Jones, G. & Deisseroth, K. (2013) Recent advances in optogenetics and pharmacogenetics. *Brain Res.*, 1511, 1-5.
- Barch, D.M., Treadway, M.T. & Schoen, N. (2014) Effort, anhedonia, and function in schizophrenia: Reduced effort allocation predicts amotivation and functional impairment. *J. Abno. Psychol.*, 123, 387-397.
- Bechara, A., Dolan, S., Denburg, N., Hindes, A., Anderson, S.W. & Nathan, P.E. (2001) Decision-making deficits, linked to a dysfunctional ventromedial prefrontal cortex, revealed in alcohol and stimulant abusers. *Neuropsychologia*, 39, 376-389.
- Bradfield, L.A., Dezfouli, A., van Holstein, M., Chieng, B. & Balleine, B.W. (2015) Medial Orbitofrontal Cortex Mediates Outcome Retrieval in Partially Observable Task Situations. *Neuron*, 88, 1268-1280.
- Brown, V.J. & Bowman, E.M. (2002) Rodent models of prefrontal function. *Trends in Neurosci.*, 25, 340-343.
- Burton, A.C., Kashtelyan, V., Bryden, D.W. & Roesch, M.R. (2014) Increased firing to cues that predict low-value reward in the medial orbitofrontal cortex. *Cerebral Cortex*, 24(12), 3310-3321.

- Carreno, F.R., Donegan, J.J., Boley, A.M., Shah, A., DeGuzman, M., Frazer, A. & Lodge, D.J. (2015) Activation of a ventral hippocampus–medial prefrontal cortex pathway is both necessary and sufficient for an antidepressant response to ketamine. *Mol. Psychiatry*, doi:10.1038/mp.2015.176
- Cavedini, P., Riboldi, G., Keller, R., D’Annunzi, A. & Bellodi, L. (2002) Frontal lobe dysfunction in pathological gambling patients. *Biol Psychiatry*, 51, 334-341.
- Fuchs, R.A., Evans, K.A., Parker, M.P. & See, R.E. (2004) Differential involvement of orbitofrontal cortex subregions in conditioned cue-induced and cocaine-primed reinstatement of cocaine seeking in rats. *J. Neurosci.*, 24, 6600-6610.
- Goodman, L.A. (1954) Kolmogorov-Smirnov tests for psychological research. *Psychol. Bulletin*, 51, 160-168.
- Gourley, S.L., Lee, A.S., Howell, J.L., Pittenger, C. & Taylor, J.R. (2010) Dissociable regulation of instrumental action within mouse prefrontal cortex. *Eur. J. Neurosci.*, 32(10), 1726-1734.
- Gourley, S.L., Zimmermann, K.S., Allen, A.G. & Taylor, J.R. (2016) The medial orbitofrontal cortex regulates sensitivity to outcome value. *J. Neurosci.*, 36, 4600-4613.
- Guettier, J.M., Gautam, D., Scarselli, M., de Azua, I.R., Li, J.H., Rosemond, E., ... & Wess, J. (2009) A chemical-genetic approach to study G protein regulation of  $\beta$  cell function in vivo. *Proc. Nat. Acad. Sci.*, 106, 19197-19202.
- Gunaydin, L.A., Yizhar, O., Berndt, A., Sohal, V.S., Deisseroth, K. & Hegemann, P. (2010) Ultrafast optogenetic control. *Nat. Neurosci.*, 13, 387-392.
- Kong, D., Tong, Q., Ye, C., Koda, S., Fuller, P.M., Krashes, M.J., ... & Lowell, B.B. (2012) GABAergic RIP-Cre neurons in the arcuate nucleus selectively regulate energy expenditure. *Cell*, 151, 645-657.

- Krashes, M.J., Koda, S., Ye, C., Rogan, S.C., Adams, A.C., Cusher, D.S., ... & Lowell, B.B. (2011) Rapid, reversible activation of AgRP neurons drives feeding behavior in mice. *J. Clin. Invest.*, 121, 1424-1428.
- Malkuz, D.C., Yenke, I., Rotella, F.M., Banakos, T., Olsson, K., Dindyal, T., ... & Bodnar, R.J. (2015) Dopamine receptor signaling in the medial orbital frontal cortex and the acquisition and expression of fructose-conditioned flavor preferences in rats. *Brain Res.*, 1596, 116-125.
- Mar, A.C., Walker, A.L., Theobald, D.E., Eagle, D.M. & Robbins, T.W. (2011) Dissociable effects of lesions to orbitofrontal cortex subregions on impulsive choice in the rat. *J. Neurosci.*, 31, 6398-6404.
- Massey Jr, F.J. (1951) The Kolmogorov-Smirnov test for goodness of fit. *J. Amer. Stat. Assoc.*, 46, 68-78.
- Nakajima, K.I., Cui, Z., Li, C., Meister, J., Cui, Y., Fu, O., ... & Wess, J. (2016) Gs-coupled GPCR signalling in AgRP neurons triggers sustained increase in food intake. *Nat. Commun.*, 7, 10268.
- Nichol, C.D. & Roth, B.L. (2009) Engineered G-protein coupled receptors are powerful tools to investigate biological processes and behaviors. *Front. Mol. Neurosci.*, 2, 16.
- Noonan, M.P., Walton, M.E., Behrens, T.E.J., Sallet, J., Buckley, M.J. & Rushworth, M.F.S. (2010) Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proc. Nat. Acad. Sci.*, 107, 20547-20552.
- Parnaudeau, S., O'Neill, P.K., Bolkan, S.S., Ward, R.D., Abbas, A.I., Roth, B.L., ... & Kellendonk, C. (2013) Inhibition of mediodorsal thalamus disrupts thalamofrontal connectivity and cognition. *Neuron*, 77, 1151-1162.

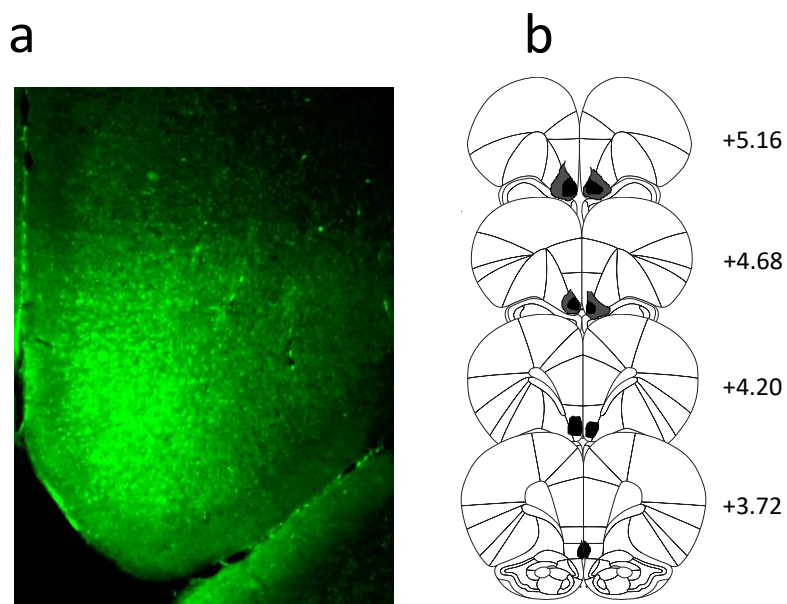
- Parnaudeau, S., Taylor, K., Bolkan, S.S., Ward, R.D., Balsam, P.D. & Kellendonk, C. (2015) Mediodorsal thalamus hypofunction impairs flexible goal-directed behaviour. *Biol. Psychiatry*, 77, 445-453.
- Paxinos, G. & Watson, C. (2007) The rat brain in stereotaxic coordinates (6<sup>th</sup> ed). Academic Press, London.
- Pei, Y., Rogan, S.C., Yan, F. & Roth, B.L. (2008) Engineered GPCRs as tools to modulate signal transduction. *Physiology*, 23, 313-321.
- Robbins, T. (2002) The 5-choice serial reaction time task: behavioural pharmacology and functional neurochemistry. *Psychopharmacology*, 163, 362-380.
- Roesch, M.R. & Olson, C.R. (2005) Neuronal activity in primate orbitofrontal cortex reflects the value of time. *J. Neurophysiol.*, 94, 2457-2471.
- Roth, B.L. (2016) DREADDs for neuroscientists. *Neuron*, 89, 683-694.
- Rudebeck, P.H. & Murray, E.A. (2014) The orbitofrontal oracle: Cortical mechanisms for the prediction and evaluation of specific behavioral outcomes. *Neuron*, 84, 1143-1156.
- Rushworth, M.F., Noonan, M.P., Boorman, E.D., Walton, M.E. & Behrens, T.E. (2011) Frontal cortex and reward-guided learning and decision-making. *Neuron*, 70, 1054-1069.
- Rushworth, M.F., Kolling, N., Sallet, J. & Mars, R.B. (2012) Valuation and decision-making in frontal cortex: One or many serial or parallel systems? *Curr. Opin. Neurobiol.*, 22(6), 946-955.
- Sachdev, P.S. & Malhi, G.S. (2005) Obsessive-compulsive behaviour: A disorder of decision-making. *Aust. & New Zeal. J. Psychiatry*, 39, 757-763.

- Stopper, C.M., Green, E.B. & Floresco, S.B. (2014) Selective involvement by the medial orbitofrontal cortex in biasing risky, but not impulsive, choice. *Cerebral Cortex*, 24(1), 154-162.
- Walton, M.E., Behrens, T.E., Noonan, M.P. & Rushworth, M.F. (2011) Giving credit where credit is due: Orbitofrontal cortex and valuation in an uncertain world. *Annals New York Acad. Sci.*, 1239(1), 14-24.
- Waltz, J.A., Frank, M.J., Robinson, B.M. & Gold, J.M. (2007) Selective reinforcement learning deficits in schizophrenia support predictions from computational models of striatal-cortical dysfunction. *Biol. Psychiatry*, 62, 756-764.
- Ward, R.D., Winiger, V., Higa, K.K., Kahn, J.B., Kandel, E.R., Balsam, P.D. & Simpson, E.H. (2015a) The Impact of Motivation on Cognitive Performance in an Animal Model of the Negative and Cognitive Symptoms of Schizophrenia. *Behav. Neurosci.*, 129(3), 292-299.
- Ward, R.D., Winiger, V., Kandel, E.R., Balsam, P.D. & Simpson, E.H. (2015b) Orbitofrontal cortex mediates the differential impact of signaled-reward probability on discrimination accuracy. *Front. Neurosci.*, 9, 230.  
doi:10.3389/fnins.2015.00230
- Yau, J.O.Y. & McNally, G.P. (2015) Pharmacogenetic excitation of dorsomedial prefrontal cortex restores fear prediction error. *J. Neurosci*, 35, 74-83.
- Zald, D.H., McHugo, M., Ray, K.L., Glahn, D.C., Eickhoff, S.B. & Laird, A.R. (2012) Meta-analytic connectivity modeling reveals differential functional connectivity of the medial and lateral orbitofrontal cortex. *Cerebral Cortex*, 24(1), 232-248.

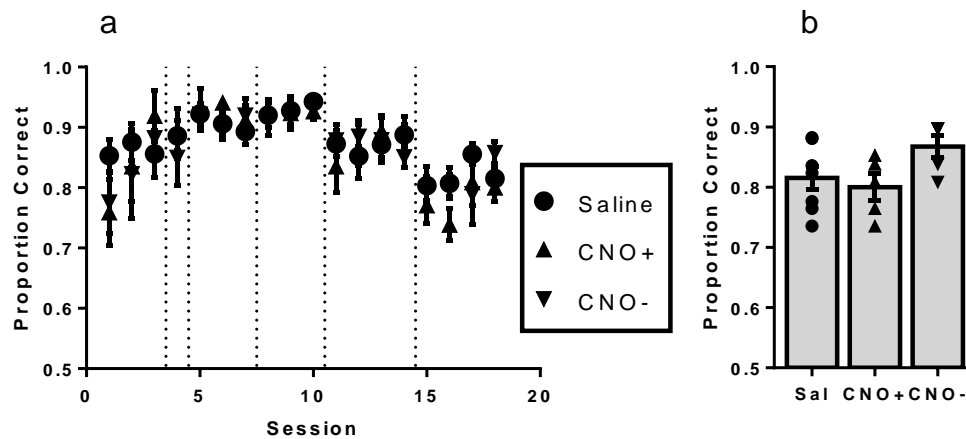


*Figure 1.* Schematic of the SPSA task. The houselight state (on or off) signals whether the current trial has a high or low reward probability for making a correct response. The cue light then signals which lever will be rewarded at the choice point. On high probability trials, pressing the correct lever always delivers reward. On low probability trials, reward receipt for a correct response is unlikely (Adapted from Ward et al., 2015a). Animals were injected with CNO or saline 30 minutes before being tested on the task once daily for 20 sessions (after Ward et al. 2015a).

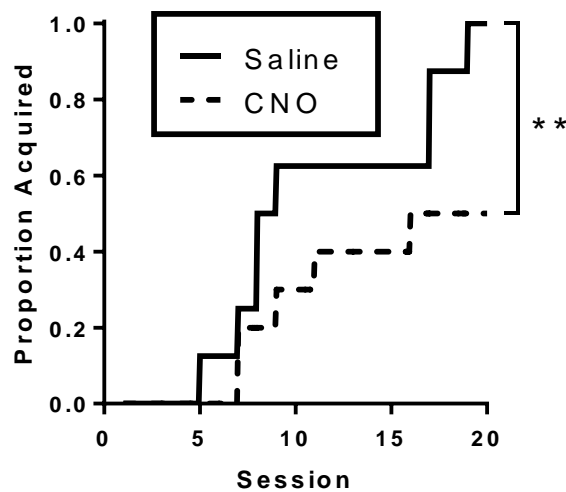




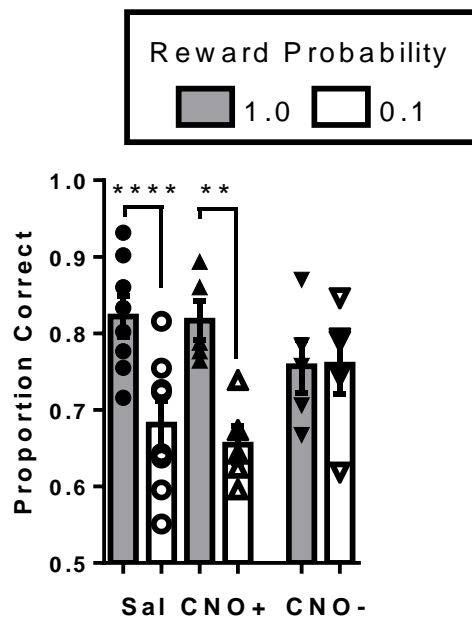
*Figure 2.* A. Representative example of viral expression in orbitofrontal cortex. B. Schematic representation of the injection location and minimal (black) and maximal (grey) spread of hM3D(G<sub>q</sub>) viral expression. All injections were bilateral. Numbers indicate relative distance from bregma according to Paxinos and Watson (2007).



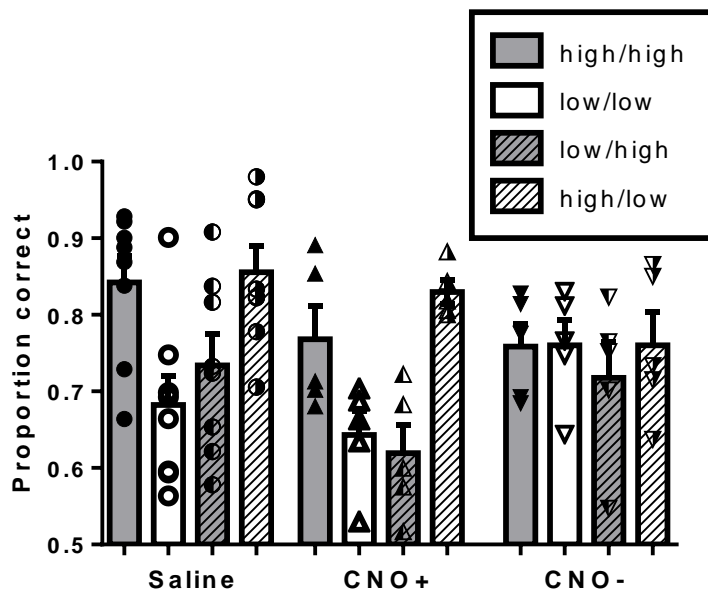
*Figure 3. A: Average proportion correct as a function of session number across the six stages of choice training and decreasing cue duration. Dashed lines indicate a change in training stage, which progressed in the following order: 50% choice, 75% choice, 100% choice, 100% choice without correction, 100% choice with 2s cue duration, 100% choice with 1s cue duration. Saline group averages are indicated by circles, whereas CNO group averages are indicated by triangles. B: Average proportion correct for Saline, CNO+ and CNO- groups on the final training session before beginning the experimental manipulation. In both panels, error bars indicate standard error of mean.*



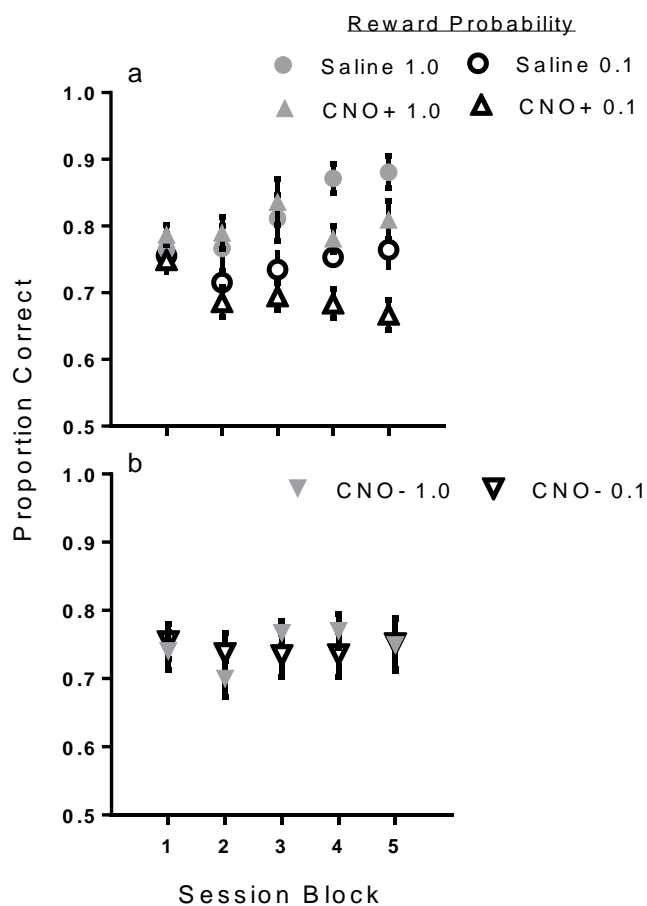
*Figure 4.* Cumulative distribution plot for SPSA acquisition. The proportion of animals that met acquisition criteria (defined as proportion correct being significantly greater on high than on low reward probability trials over three consecutive sessions) is shown on the Y-axis as a function of session number on the X-axis. Acquisition is indicated by a solid line for Saline controls and a dashed line for CNO animals. Saline  $N = 8$ , CNO  $N = 10$ . \*\*  $p < .01$



*Figure 5.* Average proportion correct on high and low probability trials. For rats that acquired (Sal and CNO+), scores come from data averaged over the three sessions which met the acquisition criterion (significantly greater proportion correct on high versus low reward probability trials over three consecutive sessions). For rats that did not acquire (CNO-) scores were averaged over the final three experimental sessions. Data points indicate data from individual rats. Error bars indicate standard error of mean. \*\*\*\*  $p < .0001$ , \*\*  $p < .01$



*Figure 6.* Average proportion correct on the current trial as a function of reward probability on the previous trial for Saline, CNO+, and CNO- rats. The first designation in each legend pair indicates the reward probability on the current trial, the second indicates the reward probability on the previous trial. Thus high/high indicates a high reward probability trial preceded by a high reward probability trial. Data points indicate data from individual rats.



*Figure 7.* A. Proportion correct on high and low reward probability trials across testing for Saline and CNO rats that met acquisition criteria. B. Proportion correct on high and low reward probability trials across testing for CNO rats that did not meet acquisition criteria.